



Db	361	GQMARGAGGGLSGYLRRPPRYMPHSFAAG	391	A; Molecule type: DNA
				A; Residues: 1-393 <COL>
A; Cross-references: GB:Z95390; GB:AL123456; NID:93261766; PIDN:CA08702.1; PID:e316074; f	C; Species: Mycobacterium tuberculosis	C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999	C; Accession: B70741	R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Shelton, S.; Squares, S.
A; Authors: Soares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.	A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome	A; Reference number: A70500; MUID:98295987; PMID:9634230	A; Accession: B70741	A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA	A; Cross-references: GB:Z75555; GB:AL123456; NID:93261608; PIDN:CAA99966.1; PID:e250360; f	A; Experimental source: strain H37RV	A; GenBank:	A; Gene: PPE
Query Match	85.0%; Score 1656.5; DB 2; Length 396;	Best Local Similarity 85.1%; Pred. No. 6.e-91;	Matches 388; Conservative 18; Mismatches 34; Indels 7; Gaps 3;	Qy 1 MVDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSAASAFOQSYVNGLTIVGSWIG 60
Db	1 MVDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSAASAFOQSYVNGLTIVGSWIG 60	Qy 61 SSAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120	Db 61 SSAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120	Db 61 SSAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120
Qy	121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEEAPEMTSAGG 180	Db 121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180	Qy 121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180	Db 121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180
Db	121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180	Db 121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180	Qy 121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180	Db 121 LIAATNLGQNTPAITAVNEAEYGENWAQDAAMFGYAAATAATLIPFEEAPEMTSAGG 180
Qy	181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240	Db 181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240	Qy 181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240	Db 181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240
Db	181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240	Db 181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240	Qy 181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240	Db 181 LLEQAAVEEASDTAAANOLMNVPQALQLAQOPTGTTPSSKLGIGWKTSPHRSPISN 240
Qy	241 MVSMANNHNSMTNSGVSMNTLSSMLKGPSAPAAQVOTAONGVRAMSS---TGSSL 296	Db 241 MVSMANNHNSMTNSGVSMNTLSSMLKGPSAPAAQVOTAONGVRAMSS---TGSSL 296	Qy 241 MVSMANNHNSMTNSGVSMNTLSSMLKGPSAPAAQVOTAONGVRAMSS---TGSSL 296	Db 241 MVSMANNHNSMTNSGVSMNTLSSMLKGPSAPAAQVOTAONGVRAMSS---TGSSL 296
Db	241 MVSMANNHNSMTNSGVSMNTLSSMLKGPSAPAAQVOTAONGVRAMSS---TGSSL 296	Db 241 MVSMANNHNSMTNSGVSMNTLSSMLKGPSAPAAQVOTAONGVRAMSS---TGSSL 296	Qy 241 VSSIANNHNSMMGCVSMNTLSSMLKGSLAP-AAAQAVETAAENGWMASSLGSQSSL 299	Db 241 VSSIANNHNSMMGCVSMNTLSSMLKGSLAP-AAAQAVETAAENGWMASSLGSQSSL 299
Qy	297 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAERGPQMG 356	Db 297 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAERGPQMG 356	Qy 297 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAERGPQMG 356	Db 297 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAERGPQMG 356
Db	300 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAQTAQPHMLG 359	Db 300 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAQTAQPHMLG 359	Qy 300 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAQTAQPHMLG 359	Db 300 GSSGLGGYGAANLGRAASYGSLSVQPAWAANQAVTPARALPLTSLTSAAQTAQPHMLG 359
Qy	357 GLPVQGMGRAG- GGLSEVLRLVPRPPVMPHSPAAAG 391	Db 357 GLPVQGMGRAG- GGLSEVLRLVPRPPVMPHSPAAAG 391	Qy 357 GLPVQGMGRAG- GGLSEVLRLVPRPPVMPHSPAAAG 391	Db 357 GLPVQGMGRAG- GGLSEVLRLVPRPPVMPHSPAAAG 391
Db	360 GLPQLTNSGGGGVSNALRMPRATVMPRVPAAG 396	Db 360 GLPQLTNSGGGGVSNALRMPRATVMPRVPAAG 396	Qy 360 GLPQLTNSGGGGVSNALRMPRATVMPRVPAAG 396	Db 360 GLPQLTNSGGGGVSNALRMPRATVMPRVPAAG 396
RESULT 3				
C; Species: Mycobacterium tuberculosis	C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999	C; Accession: C70568	R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Shelton, S.; Squares, S.	A; Authors: Soares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome	A; Reference number: A70500; MUID:98295987; PMID:9634230	A; Accession: C70568	A; Status: preliminary; nucleic acid sequence not shown; translation not shown	A; Gene: PPE
A; Molecule type: DNA	A; Cross-references: GB:AL022021; GB:AU123456; NID:93250699; PIDN:CPA17711.1; PID:e125460; f	A; Experimental source: strain H37RV	A; GenBank:	A; Gene: PPE
Query Match	40.0%; Score 779.5; DB 2; Length 393;	Best Local Similarity 43.7%; Pred. No. 4.3e-39;	Matches 179; Conservative 37; Mismatches 129; Indels 37; Gaps 10;	Qy 1 VDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSAASAFOQSYVNGLTIVGSWIG 60
Db	1 VDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSAASAFOQSYVNGLTIVGSWIG 60	Qy 62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 121	Db 62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 121	Db 62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 121
Qy	62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120	Db 62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120	Qy 62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120	Db 62 SAGLMVAAASPYVAMSVTQAGELTAQVRVAAAAYETAYGLTVPPPVIAENRAELMI 120

Qy	122 IATNLGONTPAIAVNEAFYGENWAQDAAMFGEYAATAATATLPEEAPEMTSAGGL 181	Db	353 AAEAPGALFGEMLSSLAGLAGTAVRSGAGAAARV 388
Db	121 ISTNVFGQTSAARRAAEAYGENWAQDAAMFGEYAATAATATLPEEAPEMTSAGGL 179		
Qy	182 LEQAAVEPASDAA--NOLMNNVPQLQQLAQOPTCTPSSKLGLINK---- 229		
Db	180 GTQAAAVATAAGTAQSTLTEMITGLPNALQSLSPLQS-NGPLSMNWWLFLGTTPNPFPT 238		
Qy	230 -----TSPHRSPISNVMSMANNHMMNTNSGVSMTNLSSMLKGAPAAAQAQVOTA 282		
Db	239 SISALITDQPYASFFYNTTEGLPYFSGIMGNNTQSAKTL-GHIGSAPAAVA----A 292		
Qy	283 QNGYRAMSSLGSSLGSSGGVAAANLGRASAISGVSYPOQAWA-ANCAVTPAARALPLT 341		
Db	293 GDAKGLPLGGMIG---GGPYAAGLNDASVGKLYSPVWNGPLGSVTPGAAPLVIS 348		
Qy	342 SLTSAERGPQGMGLGPQGMGARAGCCILSGYLRVPPRYMPHSAAAG 391		
Db	349 TVSAAPEAPGSLIGLPL---AGAGAGAGB-RYGFPTWMARPFAG 393		
	RESULT 5		
A70932	probable PPE protein - Mycobacterium tuberculosis (strain H37RV)		
C;Species: Mycobacterium tuberculosis			
C;Accession: A70932			
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.			
Qy	230 -----TSPHRSPISNVMSMANNHMMNTNSGVSMTNLSSMLKGAPAAAQAQVOTA 282		
Db	239 SISALITDQPYASFFYNTTEGLPYFSGIMGNNTQSAKTL-GHIGSAPAAVA----A 292		
Qy	283 QNGYRAMSSLGSSLGSSGGVAAANLGRASAISGVSYPOQAWA-ANCAVTPAARALPLT 341		
Db	293 GDAKGLPLGGMIG---GGPYAAGLNDASVGKLYSPVWNGPLGSVTPGAAPLVIS 348		
Qy	342 SLTSAERGPQGMGLGPQGMGARAGCCILSGYLRVPPRYMPHSAAAG 391		
Db	349 TVSAAPEAPGSLIGLPL---AGAGAGAGB-RYGFPTWMARPFAG 393		
	A;Cross-references: GB:AL022021; GB:AL123456; NID:93250699; PID:CAA17728_1; PID:e125461		
A;Experimental source: strain H37RV			
C;Genetics:			
A;Gene: PPE			
	RESULT 6		
H70931	probable PPE protein - Mycobacterium tuberculosis (strain H37RV)		
C;Species: Mycobacterium tuberculosis			
C;Accession: H70931			
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.			
Nature 393, 537-544, 1998			
A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.			
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome			
A;Reference number: A70500; PMID:9634230			
A;Accession: H70931			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-403 <COL>			
A;Cross-references: GB:AL022021; GB:AL123456; NID:93250699; PID:CAA17728_1; PID:e125461			
A;Experimental source: strain H37RV			
C;Genetics:			
A;Gene: PPE			
	RESULT 7		
B70931	probable PPE protein - Mycobacterium tuberculosis (strain H37RV)		
C;Species: Mycobacterium tuberculosis			
C;Accession: B70931			
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.			
Nature 393, 537-544, 1998			
A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.			
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome			
A;Reference number: A70500; PMID:9634230			
A;Accession: B70931			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-403 <COL>			
A;Cross-references: GB:AL022021; GB:AL123456; NID:93250699; PID:CAA17728_1; PID:e125461			
A;Experimental source: strain H37RV			
C;Genetics:			
A;Gene: PPE			
	RESULT 8		
Qy	2 VDFGALPPBINSARMYAGGGSASLVALAQNWDSVASDLFSAASFQSVWGLTVGSWIGS 61	Db	223 KLGGLWKTSPHERSPISNNVMSMNTNSGYSMTNTLS---SMLKFAPAAAQAV 278
Db	1 MDFFGALPPBINSCRMAYAGGGSGPLAAAAWDLAAEYSSAASYGTTIEGLTVAPMGIP 60		
Qy	62 SAGLMVAAASPYAAMSVTAGQAELTAQVRVAAAAYETAYGLTVPPPVIANRAELML 121		
Db	61 SSITMAAAAPYYAWISPTAGQAEQAAQKIAQVYETATAATVPPVIANRALIMSL 120		
Qy	122 IATNLGONTPAIAVNEAFYGENWAQDAAMFGEYAATAATATLPEEAPEMTSAGGL 181		
Db	121 VATNLGONTPAIAVNEAFYGENWAQDAAMFGEYAATAATATLPEEAPEMTSAGGL 179		
Qy	182 LEQAAVEPASDAA-----NQLMNNVPQALQQLAQPTQGTPSSKLGLWKTVSPH 234		
Db	180 AACSAVVAAGAAASSDITAQISQLISLSPSTIQLSA-TPATATSASAG-WDTV-- 232		
Qy	235 RSPSNMVSMANNHMSMNTSGVS-----MTNTLSSMLKGFPAAAQVQTAQNGYRAM 289		
Db	233 LQSTITLNLNTGPYSTICLGATFGQWLGLAQNAQNPQVAAAGPKAAAGALSP 292		
Qy	290 SSL-GSSLSS-SGIJGGYAAANLGRASAISGVSYPOQAWAANQVTPAARALPLTSLSA- 346		
Db	293 APLRGYIGDITPLGGGATGGIARAIYVGSLSPQGWAAAPMVRATASVLPGTGAAPAL 352		
Qy	347 AERSPGQMIQGLPVQGMGARAGGGL----SGYLRV 377		

Query Match Similarity 37.5%; Score 731; DB 2; Length 421;							
Best Local Similarity 41.1%; Pred. No. 3-36; Matches 174; Conservative 55; Mismatches 160; Indels 34; Gaps 8;							
A,Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.	A,Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome	A,Reference number: A70500; MUID:91250099; PMID:1963230	A,Accession: B70931	A,Status: preliminary; nucleic acid sequence not shown; translation not shown	A,Molecule type: DNA	A,Residues: 1-423 <COL>	A,Cross references: GB:AL022021; GB:AU123456; NID:91300999; PIDN:CAA17722.1; PMID:e125461
A,Experimental source: strain H37Rv	C,Genetics:						
Query Match Score 37.5%; Score 731; DB 2; Length 423;	Best Local Similarity 41.1%; Pred. No. 3-36; Matches 174; Conservative 55; Mismatches 160; Indels 34; Gaps 8;	Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;					
Qy 2 VDFGALPPBINSARMAMYAGGSASIVAAQMWDSVASDLESAASFQSVWVGTLVGSWIGS 61	Ddb 1 MDPFULLPBPBINSARMYTGPPGMLAATAWDGLAVELHATAYGASLISALT-GAGNSGP 59						
Qy 62 SAGIMVAAAASPYYWMSVTAQGAELTAACVRVAAAYETAYGLTVPPIVIAENRAELMIL 121	Ddb 60 SSSTMASAAAPYIWWMSATAVHELAGOARLAAYDAAPATVPPPIVIAANRAQLMVL 119						
Qy 122 IATNLIGONTPAIAVNEAEYGEMWAQDAAMFGYAAATATATATLPLFEEAPEMTSAGGL 181	Ddb 120 IATNIFGQTTPAINTTEAOYMEWAQDAAMYGSSATA-SRMTAATTEPPPTNHGQL 178						
Qy 182 LEQAAVEEASDTAAAN----QLMMNVQPQLQLAQPT---QGTTP-----S 221	Ddb 179 GAQSSAVQATAAAGGNIQSAFFQLSSAVPRAQGIALPTAQOSASATPQWVTDLGRLS 238						
Qy 222 SKUGGLWIKTVSPHSRISPISNMVSMANNMSMTNSGSMTNTLSSMLKGPAAPAAAQVOTA 281	Ddb 239 TFLGG---AVTGPVTFP-----GVLPSSCPVYLGIQSVL-----V 271						
Qy 282 AQNQVRANSSLGS-----SLGSGSLGGG-VAANLGRASAQSGLS 319	Ddb 272 TQNGQVSALLGKGGKPTIGALAPLAFAFLHTPILQPLGSEGLGGSVSAGIGRAGLVKLS 331						
Qy 320 VPOQWAANQAVTAARALPLTSLS--AAERPGQOMGQGLPVQGARGGGLGSVLR 376	Ddb 332 VPQGWTVAAPIEIPSAAALQATRLLAAPTAATGAGALGMALSGLGRAAAGSTG-- 388						
Qy 377 VPPRPYMPHSPPAAG 391	Ddb 389 --HPIGSAAAAPAVG 400						
RESULT 8							
HB1056	PPB-f-family protein [imported] - Mycobacterium leprae						
C,Species: Mycobacterium leprae	C,Accession: B87056						
C,Title: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001	C,Accession: B87056						
C,Authors: Eigner, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holroyd, M.A.; Rutherford, K.M.; Davies, R.M.; Devlin, K.; Dutchoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, M.A.; Rutherford, K.M.;	C,Accession: B87056						
C,Title: Massive gene decay in the leprosy bacillus.	C,Accession: B87056						
C,Status: Preliminary	C,Accession: B87056						
C,Molecule type: DNA	C,Accession: B87056						
C,Residues: 1-421 <STO>	C,Accession: B87056						
C,Genetics: ML1182	C,Accession: B87056						

RESULT 11  
 Db 180 AGQQAAATVSSTVEPLATAAVPQLCOLSTSLLPWSALQWIAENLIGLTDPDNMTIV 239  
 Qy 225 -----GGIWKTVSPRSPESNMVSANNAMSMNSGVSMINTLSSMKGAFAPAAA 275  
 Db 240 RLLGTSYFPEGL-----LQEFAASLQQAIPGTGAG--DSSSSLDWSWPTEA 287  
 Qy 276 QAVOTAAQNGVRAMSSL-GSSLS-----SGUGGYVAANLGRAASVGSL 319  
 Db 288 -----GPRAPSVAAGGGAVQPYWWALDRSITGSVISAAALGKGSAGSL 338  
 Qy 320 VPOAWAAAQVTPAARALP---LTSLTSAAERGPOMLGLPVONGARAGGGLSVLR 376  
 Db 339 VPPDKWAARRWANDAAWRDLPDDVTALRGTAENA--LURGFPMASAGQSTGGF--VHK 393  
 Qy 377 VPPRPPYMHPSPAAG 391  
 Db 394 YGFLALMORPPFAG 408

RESULT 10  
 B70625  
 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Accession: B70625  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A;Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID: 38295387; PMID: 9634220  
 A;Accession: P70625  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Cross-references: 1-413 <COL>  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: PPE

Query Match 36.0%; Score 702; DB 2; Length 413;  
 Best Local Similarity 39.1%; Prec. No. 1.7e-34;  
 Matches 168; Conservative 61; Mismatches 145; Indels 56; Gaps 7;

Qy 1 MVDFGALPPETNSARMYAGPGSASLIVAAQMWDSVASYDLSFSAAGAFQSYYWGLTVGSGWIG 60  
 Db 1 MLDFDALPPETNSALMYAGPGSASLIVAAQMWDSVASYDLSFSAAGAFQSYYWGLTVGSGWIG 60  
 Qy 61 SSAGLMVAASPYWAMSVTAGOABLTAQVRVAAAYETAYGHTVPPPYIAENRAELMI 120  
 Db 61 SSAASMVAAATPQAVWLRSFGAOFAQGSQAVAAASAAFFATVPPPEIAANRAELMA 120  
 Qy 121 LIATNLGQNPAIAVNEABGEWAQDAAMFGTAATAATLLPFEAPEITSAGG 180  
 Db 121 LLATNLGQNPAIAAATQAYEAATQADAAAMYGYAGASAATQSPNPNAAQTINPAG 179  
 Qy 181 LLEOAIAVEEASDAAANOLMNNYQOALQOQAOPFTGTPSSKLQGLWKTSPHRSPISN 240  
 Db 180 LASQASVGAVASGNAQALTDIFRL-----FGLSGIFPNEPPWLTDLGK 226  
 Qy 241 MVSMANHMSMTNSGYSMTLSSMLKGFPAAA---AAQAVOT-----  
 Db 227 ALGLTGHWTSDGSGLIVGGVLGDVQGTGSAEELDASYAMDTEFGKWSPARLMVTFKD 286  
 Qy 281 -----AQNQVBRMSSLOSSCGEGVAANLGRASVGSLSUPQA 323  
 Db 287 YFGLAHDLPKWASEGAKAAGEAAKALPAAPVPSGL-SGVAGAVGQAAVGCGKVPAY 345  
 Qy 324 WAANGAAVTPAARALPLTSLTSAAERGPOMLGGLPVONGARAGGGLSGVL--RVPFRD 381  
 Db 346 WTATTFAASAVLA-SNGLGAIAAEGSTHAFGMP-L--MGSGJGRAFINFEAARYGFRP 403  
 Qy 382 YVMPHSPAAAG 391  
 Db 404 TVIAQPFPAGG 413

RESULT 12  
 C70931  
 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Accession: C70931  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A;Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

Query Match 36.0%; Score 702; DB 2; Length 391;  
 Best Local Similarity 42.3%; Prec. No. 1.e-34;  
 Matches 166; Conservative 62; Mismatches 150; Indels 14; Gaps 7;

Qy 62 SAGIMVAASPYWAMSVTAGOELTAAQVRVAAAYETAYGHTVPPPYIAENRAELMI 121  
 Db 61 ASHANVAAAPQIYLWLTAAAHAGSQAMSAAAEEAYANTVPEPEVAANRALLAAL 120  
 Qy 122 IATNLGQNPAIAVNEABGEWAQDAAMFGTAATAATLLPFEAPEITSAGG 181  
 Db 121 VATNLGQNPAIAAATQAYEAATQADAAAMYGYAGASAATQSPNPNAAQTINPAG 179  
 Qy 182 LEQAAVEEASDAAANQ---LMNNPQALQOQAQPTGTTPSKLGWMKTVSPHS 236  
 Db 180 AAQAAVGSAAATAVNVSVADLISSPNAYSVLSDSGTSGIADIALLA 239  
 Qy 237 P--ISNNMYSMANNMS-MNNSGYSMTLSSMLKGFPAPAALQAVTAQVRAANSSLG 293  
 Db 240 TPPFANILINASVNTAAWYNAAPTAFLANALNSGAPVATEAEG--AASAAA 296  
 Qy 294 SSLGSSGCGGGVAANLGRAASVSSLSPQOWAAANQATTAARALPLTSLSAAERGPQQ 353  
 Db 297 AGLADSVTFPLGASLGTATVGLSRSPAAWSTAATAGTAGTAALEGSGWTVAEEA-GP 355  
 Qy 354 MLGGLPVGCMGARAGGHLGVRLVPPRYVMP 385  
 Db 356 VTGMP-GRASAAGTGAXAGPRYGEKTPVMP 386.

C:Genetics: C:Experimental source: strain H37Rv							
A:Genes: PPE							
Query	Match	Score 35.3%	Length 463;				
Best Local Similarity	42.8%	Pred. No. 1.3e-33;					
Matches	17;	Conservative	48; N mismatches 136;	Indels 49;	Gaps 11;		
Qy	2 VDFGALPPEINSARNMYAGPGSASLYAAQNMWDSASLPSAASFQSVVWGLTVGSWIGS 61						
Db	1 MDFGALPPEINSGNMYAGPGSPMIAAAAADGIALELTQSTAAYGSVISVL-T-GVNSGQ 59						
Qy	62 SAGIMVAAAAPYVWMSVTAQGAELTTAAQVRAAAYEATAGLVTPPPVTAERAEMLI 121						
Db	60 SSGTMAAAAAPYVWMSVTAALAREAAASAAAYEAAPATVPPVVAATRAELAVI 119						
Qy	122 IATNLIGONTPAIAYNEAEYGMWQDAALAMFGYAATAATATATLPPPEAPMTNSAGGL 181						
Db	120 AATNFQNGIAALAAEYAYEMAQDAMGYAGSSVAT-QIVTPFAPPPTNAGL 178						
Qy	182 LEQQAAVEEASDTAAQNLMMNPQALQIAOPPGTTPSKLGGLMKTVS--PHRSPI- 238						
Db	179 ATGGVAAQVGASBAGN-ARSIVLSEVLFIA--TAGTNTNKTVTASLMNAVTCYVASSV 235						
Qy	239 ----SNNTMAMNMMSMTNSGVNTINTLSSMKGFAPMAAQAQVTAQNQTRIA 288						
Db	236 NSMLGIGFAKSMMVLPANDITVISTIETGMYQFKENPVTPFPNPLIPK----- 283						
Qy	289 MSSGSSSLG----SSGLGS--GVAANILGRAASVGSLSVPQAMAANAOAVTPAARALPL 340						
Db	284 -SAGAGGURSAASSGGSTAPASAGASQAGSVCAGSVPSSMAARTPAIRTYAVFSS 342						
Qy	341 TSLSI--AAERGPQGML-----GGLPVGQMGSARAGGLSGTLRV 377						
Db	343 TGLQAVPAAAISEGSLLSQMALASVAGGLGAAARATGGFLGGGRV 389						
RESULT 13							
B70932	probable PPE protein - Mycobacterium tuberculosis (strain H37RV)						
C:Species: Mycobacterium tuberculosis							
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999							
C:Accession: B70932							
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.							
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome							
A:Reference number: A70930; PMID:98295387; PMID:9634320							
A:Accession: B70932							
A:Status: preliminary; nucleic acid sequence not shown; translation not shown							
A:Residues: 1-468 <COL>							
A:Cross-references: GB:AL022021; GB:AL123456; NID:93250699; PID:CAA17730.1; PID:e125462							
C:Genetics: C:Experimental source: strain H37Rv							
A:Genes: PPE							
Query	Match	Score 35.0%	Length 468;				
Best Local Similarity	40.9%	Pred. No. 2.9e-33;					
Matches	164;	Conservative	59; N mismatches 125;	Indels 53;	Gaps 9;		
Qy	2 VDFGALPPEINSARNMYAGPGSASLYAAQNMWDSASLPSAASFQSVVWGLTVGSWIGS 61						
Db	1 MDFGALPPEINSGNMYAGPGSPMIAAAAADGIALELTQSTAAYGSVISVL-T-GVNSGQ 60						

Db 290 GAASVGNVTLASVGRANSIGQLSVPPSWAAPSTRPVSAASPAGLTLIPGTDVAEHGMPG- 348  
 Qy 354 MLGGIPVGCGMARGAGGGLGSVGVLPVPPYVMHSPAG 391  
 Db 349 -VPGVPV----AAGRASGVLPFGVRLTMAHPFAG 380

## RESULT 15

G70881 probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C;Accession: G70881  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentle, S.; Hamlin, N.; Holroyd, S.  
 Randal, M.A.; Rogers, J.J.; Rutten, S.; Seeger, K.; Shelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Soares, P.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; PMID:98255987; PMID:9634230  
 A;Accession: G70881  
 A;status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-394 <COL>  
 A;Cross-references: GB:AL008967; GB:AL123456; NID:93261491; PID:CAA15564.1; PID:e117389  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: PPE

Query Match	34.3%	Score	668.5	DB	2	Length	394
Best Local Similarity	40.8%	Pred.	No.	1	.6e-32		
Matches	161;	Mismatches	156;	Indels	17;	Gaps	7;
Matches	161;	Conservative	61;				

Qy 2 VDFFGALPPIINSARMYAGPGSASILVAAQMMDSVASDLSASAASFQSYVWGLTVGSWIGS 61  
 Db 1 MDFGALPPIINSRYAGGAAPIMAATGQVNLAVLTLVSLAVELSTTASSVESPIMQETTEOWLGP 60

Qy 62 SAGIMVAAAMSPYAWMSUTAGQELTAQVNRVAAAYETAYGLTVPPPVIAENRALLML 121  
 Db 61 ASMVMVVAAPQYLAWLYTVAESAAHAAQAMASAAAFAAFAFMTPVPAEVANRALLAL 120

Qy 122 IATNLLGQNTPAIAVNBEAYGEWQAQDAAMFGYAAATAATATLIFEEADBMTSAGGL 181  
 Db 121 VATNVLGQNTPAIMATEAHYGEWQAQDAALAMYAASSA-AAGRINLPLITPSQTANMAGL 179

Qy 182 LEQAAAVEREASDTAAANO---LMNNVYQALQQLAQPTQGTTPSSKIGJNKRTVSP--H 234  
 Db 180 AGQAAVSISSAAASTVQQIGLGLISNLDNAVNGFASPLTSADAAGIGGIQDIEBLIG 239

Qy 235 RSPISNNMYSMANNMMSMTNSGVSMTNTLSMKGFAAAAQAVQTAQNGY--RAMSS 291  
 Db 240 ITFVQNAINGAVN---TTAWFMATIPAVFLGHAFALNPATVTAADAVPAAAAAG 295

Qy 292 LGSSLGSSLG-GVAAANIGRAASVGSLSVPOQWAANQAVT'FAARALPLTUTSAERG 350  
 Db 296 LAHTVTPVGVGGASLTASIGEASSVGGSLSPAGWSSTAAPMTSGTTALEGSSWAVPEEAG 355

Qy 351 PGOMIGGLPVGQMPARAGGGLGLVRPPRYWMP 365  
 Db 356 PVAAMPGM-AGISGAAKGMAYAGPRYGFKPITMP 369

Search completed: August 25, 2004, 03:13:46  
 Job time : 41 secs

This Page Blank (Top)

GenCore version 5.1.6						
Copyright (c) 1993 - 2004 Compugen Ltd.						
run on: August 25, 2004, 03:05:03 ; Search time 118 Seconds (without alignments)						
scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5						
title: US-09-724-685-107 perfect score: 1949 sequence: 1 MVDFGALPPPEINSARMYAGP.....SGVLRVPPRYPMPHSPAAAG 391						
Searched: 1017041 seqs, 315518202 residues Total number of hits satisfying chosen parameters: 1017041						
Minimum DB seq length: 0 Maximum DB seq length: 2000000000 Post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries						
SPREMBL_25:*						
1: sp_archaea:*						
2: sp_bacteria:*						
3: sp_fungi:*						
4: sp_human:*						
5: sp_invertebrate:*						
6: sp_mammal:*						
7: sp_mic:*						
8: sp_organelle:*						
9: sp_phage:*						
10: sp_plant:*						
11: sp Rodent:*						
12: sp_virus:*						
13: sp_vertebrate:*						
14: sp_unclassified:*						
15: sp_virus:*						
16: sp_bacteriopl:						
17: sp_archaeap:*						
SPREMBL :						
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
No.	Score	Query	Match	Length	DB	ID
Requirement		Description				
1	1949	005298	Q7u05	16	005298	Q7u05
2	1920.5	005298	Q7u05	16	Q7u05	Q7u05
3	1779.5	005298	Q7u05	16	Q7u05	Q7u05
4	1648	005298	Q7u05	16	Q7u05	Q7u05
5	1583	005298	Q7u05	16	Q7u05	Q7u05
6	1583	005298	Q7u05	16	Q7u05	Q7u05
7	779.5	005298	Q7u05	16	Q7u05	Q7u05
8	779.5	005298	Q7u05	16	Q7u05	Q7u05
9	754	005298	Q7u05	16	Q7u05	Q7u05
10	751	005298	Q7u05	16	Q7u05	Q7u05
11	749	005298	Q7u05	16	Q7u05	Q7u05
12	740.5	005298	Q7u05	16	Q7u05	Q7u05
13	737.8	005298	Q7u05	16	Q7u05	Q7u05
14	736	005298	Q7u05	16	Q7u05	Q7u05
15	736	005298	Q7u05	16	Q7u05	Q7u05
16	731	005298	Q7u05	16	Q7u05	Q7u05
17	727	005298	Q7u05	16	Q7u05	Q7u05
18	702	005298	Q7u05	16	Q7u05	Q7u05
19	693	005298	Q7u05	16	Q7u05	Q7u05
20	682.5	005298	Q7u05	16	Q7u05	Q7u05
21	682.5	005298	Q7u05	16	Q7u05	Q7u05
22	675	005298	Q7u05	16	Q7u05	Q7u05
23	675	005298	Q7u05	16	Q7u05	Q7u05
24	670.5	005298	Q7u05	16	Q7u05	Q7u05
25	669.5	005298	Q7u05	16	Q7u05	Q7u05
26	668.5	005298	Q7u05	16	Q7u05	Q7u05
27	667.5	005298	Q7u05	16	Q7u05	Q7u05
28	666.5	005298	Q7u05	16	Q7u05	Q7u05
29	641	005298	Q7u05	16	Q7u05	Q7u05
30	637	005298	Q7u05	16	Q7u05	Q7u05
31	635	005298	Q7u05	16	Q7u05	Q7u05
32	633.5	005298	Q7u05	16	Q7u05	Q7u05
33	629.5	005298	Q7u05	16	Q7u05	Q7u05
34	626.5	005298	Q7u05	16	Q7u05	Q7u05
35	618.5	005298	Q7u05	16	Q7u05	Q7u05
36	615.5	005298	Q7u05	16	Q7u05	Q7u05
37	615.5	005298	Q7u05	16	Q7u05	Q7u05
38	610	005298	Q7u05	16	Q7u05	Q7u05
39	604	005298	Q7u05	16	Q7u05	Q7u05
40	603	005298	Q7u05	16	Q7u05	Q7u05
41	603	005298	Q7u05	16	Q7u05	Q7u05
42	597.5	005298	Q7u05	16	Q7u05	Q7u05
43	594.5	005298	Q7u05	16	Q7u05	Q7u05
44	589	005298	Q7u05	16	Q7u05	Q7u05
45	589	005298	Q7u05	16	Q7u05	Q7u05
46	589	005298	Q7u05	16	Q7u05	Q7u05
47	589	005298	Q7u05	16	Q7u05	Q7u05
48	589	005298	Q7u05	16	Q7u05	Q7u05
49	589	005298	Q7u05	16	Q7u05	Q7u05
50	589	005298	Q7u05	16	Q7u05	Q7u05
51	589	005298	Q7u05	16	Q7u05	Q7u05
52	589	005298	Q7u05	16	Q7u05	Q7u05
53	589	005298	Q7u05	16	Q7u05	Q7u05
54	589	005298	Q7u05	16	Q7u05	Q7u05
55	589	005298	Q7u05	16	Q7u05	Q7u05
56	589	005298	Q7u05	16	Q7u05	Q7u05
57	589	005298	Q7u05	16	Q7u05	Q7u05
58	589	005298	Q7u05	16	Q7u05	Q7u05
59	589	005298	Q7u05	16	Q7u05	Q7u05
60	589	005298	Q7u05	16	Q7u05	Q7u05
61	589	005298	Q7u05	16	Q7u05	Q7u05
62	589	005298	Q7u05	16	Q7u05	Q7u05
63	589	005298	Q7u05	16	Q7u05	Q7u05
64	589	005298	Q7u05	16	Q7u05	Q7u05
65	589	005298	Q7u05	16	Q7u05	Q7u05
66	589	005298	Q7u05	16	Q7u05	Q7u05
67	589	005298	Q7u05	16	Q7u05	Q7u05
68	589	005298	Q7u05	16	Q7u05	Q7u05
69	589	005298	Q7u05	16	Q7u05	Q7u05
70	589	005298	Q7u05	16	Q7u05	Q7u05
71	589	005298	Q7u05	16	Q7u05	Q7u05
72	589	005298	Q7u05	16	Q7u05	Q7u05
73	589	005298	Q7u05	16	Q7u05	Q7u05
74	589	005298	Q7u05	16	Q7u05	Q7u05
75	589	005298	Q7u05	16	Q7u05	Q7u05
76	589	005298	Q7u05	16	Q7u05	Q7u05
77	589	005298	Q7u05	16	Q7u05	Q7u05
78	589	005298	Q7u05	16	Q7u05	Q7u05
79	589	005298	Q7u05	16	Q7u05	Q7u05
80	589	005298	Q7u05	16	Q7u05	Q7u05
81	589	005298	Q7u05	16	Q7u05	Q7u05
82	589	005298	Q7u05	16	Q7u05	Q7u05
83	589	005298	Q7u05	16	Q7u05	Q7u05
84	589	005298	Q7u05	16	Q7u05	Q7u05
85	589	005298	Q7u05	16	Q7u05	Q7u05
86	589	005298	Q7u05	16	Q7u05	Q7u05
87	589	005298	Q7u05	16	Q7u05	Q7u05
88	589	005298	Q7u05	16	Q7u05	Q7u05
89	589	005298	Q7u05	16	Q7u05	Q7u05
90	589	005298	Q7u05	16	Q7u05	Q7u05
91	589	005298	Q7u05	16	Q7u05	Q7u05
92	589	005298	Q7u05	16	Q7u05	Q7u05
93	589	005298	Q7u05	16	Q7u05	Q7u05
94	589	005298	Q7u05	16	Q7u05	Q7u05
95	589	005298	Q7u05	16	Q7u05	Q7u05
96	589	005298	Q7u05	16	Q7u05	Q7u05
97	589	005298	Q7u05	16	Q7u05	Q7u05
98	589	005298	Q7u05	16	Q7u05	Q7u05
99	589	005298	Q7u05	16	Q7u05	Q7u05
100	589	005298	Q7u05	16	Q7u05	Q7u05
101	589	005298	Q7u05	16	Q7u05	Q7u05
102	589	005298	Q7u05	16	Q7u05	Q7u05
103	589	005298	Q7u05	16	Q7u05	Q7u05
104	589	005298	Q7u05	16	Q7u05	Q7u05
105	589	005298	Q7u05	16	Q7u05	Q7u05
106	589	005298	Q7u05	16	Q7u05	Q7u05
107	589	005298	Q7u05	16	Q7u05	Q7u05
108	589	005298	Q7u05	16	Q7u05	Q7u05
109	589	005298	Q7u05	16	Q7u05	Q7u05
110	589	005298	Q7u05	16	Q7u05	Q7u05
111	589	005298	Q7u05	16	Q7u05	Q7u05
112	589	005298	Q7u05	16	Q7u05	Q7u05
113	589	005298	Q7u05	16	Q7u05	Q7u05
114	589	005298	Q7u05	16	Q7u05	Q7u05
115	589	005298	Q7u05	16	Q7u05	Q7u05
116	589	005298	Q7u05	16	Q7u05	Q7u05
117	589	005298	Q7u05	16	Q7u05	Q7u05
118	589	005298	Q7u05	16	Q7u05	Q7u05
119	589	005298	Q7u05	16	Q7u05	Q7u05
120	589	005298	Q7u05	16	Q7u05	Q7u05
121	589	005298	Q7u05	16	Q7u05	Q7u05
122	589	005298	Q7u05	16	Q7u05	Q7u05
123	589	005298	Q7u05	16	Q7u05	Q7u05
124	589	005298	Q7u05	16	Q7u05	Q7u05
125	589	005298	Q7u05	16	Q7u05	Q7u05
126	589	005298	Q7u05	16	Q7u05	Q7u05
127	589	005298	Q7u05	16	Q7u05	Q7u05
128	589	005298	Q7u05	16	Q7u05	Q7u05
129	589	005298	Q7u05	16	Q7u05	Q7u05
130	589	005298	Q7u05	16	Q7u05	Q7u05
131	589	005298	Q7u05	16	Q7u05	Q7u05
132	589	005298	Q7u05	16	Q7u05	Q7u05
133	589	005298	Q7u05	16	Q7u05	Q7u05
134	589	005298	Q7u05	16	Q7u05	Q7u05
135	589	005298	Q7u05	16	Q7u05	Q7u05
136	589	005298	Q7u05	16	Q7u05	Q7u05
137	589	005298	Q7u05	16	Q7u05	Q7u05
138	589	005298	Q7u05	16	Q7u05	Q7u05
139	589	005298	Q7u05	16	Q7u05	Q7u05
140	589	005298	Q7u05	16	Q7u05	Q7u05
141	589	005298	Q7u05	16	Q7u05	Q7u05
142	589	005298	Q7u05	16	Q7u05	Q7u05
143	589	005298	Q7u05	16	Q7u05	Q7u05
144	589	005298	Q7u05	16	Q7u05	Q7u05
145	589	005298	Q7u05	16	Q7u05	Q7u05
146	589	005298	Q7u05	16	Q7u05	Q7u05
147	589	005298	Q7u05	16	Q7u05	Q7u05
148	589	005298	Q7u05	16	Q7u05	Q7u05
149	589	005298	Q7u05	16	Q7u05	Q7u05
150	589	005298	Q7u05	16	Q7u05	Q7u05
151	589	005298	Q7u05	16	Q7u05	Q7u05
152	589	005298	Q7u05	16	Q7u05	Q7u05
153	589	005298	Q7u05	16	Q7u05	Q7u05
154	589	0052				





Db	300	GSSGIGAYAANIGRAASVGSLVPPVAAANGAVTAARALPLTSLSAQTAPGHMLG	359	
Qy	357	GLPYGQMARGAGGLSGVLRVPPRPPYMPHSPAG	391	
Db	360	GLPLGH-SVNAAGSCINNALRVPARAYAIPTPAAG	393	
				RESULTS 6
O7TWF5		PRELIMINARY;	PRT;	393 AA.
ID	Q7TWF5;			
AC				RP
DT	01-OCT-2003	(TREMBLrel. 25, Last sequence update)	SEQUENCE FROM N.A.	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	STRAIN=AF2122/97;	
RA	Garnier T., Eigmüller K., Camus J.-C., Medina N., Mansoor H.,	Mansoor H., Mansoor H.,	RC	
RA	Pryor M., Durhoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,	Camus J.-C., Monsenpe C., Simon S.,	RX	
RA	Harris B., Atkin R., Doggett J., Keating L., Wheeler P.R.,	Lacroix C., Grondin S., Mayes R., Keating L., Wheeler P.R.,	MEDLINE=22709107; PubMed=12788972;	
DE	PE family protein.		RA	
GN	PPB60 OR MB3505.		Prior M., Durhoy S., Grondin S., Mayes R., Keating L., Wheeler P.R.,	
OS	Mycobacterium bovis.		RA	
OC	Bacteria; Actinobacteria; Actinomycetales; Mycobacterium.		Parikh J., Barriel B.G., Cole S.T., Gordon S.V., Hewinson R.G.;	
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.		RT	"The complete genome sequence of Mycobacterium bovis.";
NCBI_TaxID=165;			Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).	
RN			RL	
			DR	
			EMBL; BX248340; CAD9450.1; -.	
KW	Complete proteome.		KW	
SQ	SEQUENCE 393 AA;	38572 MW;	SEQUENCE 393 AA;	41944B6E547A8AB0 CRC64;
O7TWF5		PRELIMINARY;	PRT;	393 AA.
ID	Q7TWF5;			Query Match 40.0% Score 779.5; DB 16; Length 393;
AC				Best Local Similarity 43.7%; Pred. No. 2e-36;
DT	01-OCT-2003	(TREMBLrel. 25, Created)	Matches 179; Conservative 65; Mismatches 122; Indels 37; Gaps 10;	
RA	Garnier T., Eigmüller K., Camus J.-C., Medina N., Mansoor H.,			
RA	Pryor M., Durhoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,			
RA	Harris B., Atkin R., Doggett J., Keating L., Wheeler P.R.,		Qy 2 VDFGALPPEINSARMYAGPGSASIVAAQMDVSASDLSVQELTAQVRYAAAYETAYGTIVPPVIAANRAELMIL 61	
RA	Parikh J., Barriel B.G., Cole S.T., Gordon S.V., Hewinson R.G.;		Db 1 MDFGALPPEINSARMYAGPGSASIVAAQMDVSASDLSVQELTAQVRYAAAYETAYGTIVPPVIAANRAELMIL 60	
RT	"The complete genome sequence of Mycobacterium bovis".;			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).		Qy 62 SAGLMVAALSPTVAMSTTAGQELTAQVRYAAAYETAYGTIVPPVIAANRAELMIL 121	
DR	EMBL; BX248346; CAD95692.1; -.		Db 61 ASAAAEEAVAPYAMSAAAAQQAEQAAQCARAAAAAAFAFATVPPLIAANRASIMQL 120	
KW	Complete proteome.			
SQ	SEQUENCE 393 AA;	39413 MW;	Qy 122 IATNLIGONTPATAVNEAEYGENWAQDAAMFGYAAATAATATLPEEAPEMTSGAGL 181	
			Db 121 ISTNVFGQNTSATAAAEEQGENWAQDAAMTAYAGSSASA-SVTFSTPQIANTPAQ 179	
			Qy 182 LEQAAAVEEASDPA--NQLMANNVPOQIQLQOPTQCTTPSSKLGWIK----- 229	
O7TWF5		PRELIMINARY;	Db 180 GTQAAAVATAGTAQSTLTEMITGLPNLQSLTSPLLOSS-NGPLSWIQLFLGTNPFT 238	
ID	Q7TWF5;		Qy 230 -----TSPHRSPIISNVSMANNHMNTNSCEVSMTNLKGFPAAAQQAVOTAA 282	
AC			Db 239 SISALLTDIQQPYAFFNTIEGLPYFSIGNNNNTIQAATL-GLIGSAPPAAVA----AA 292	
DT	01-OCT-2003	(TREMBLrel. 25, Last sequence update)	Qy 283 QNGYRAMSSGSSGSSGIGGYAANIGRAASGSSLSPQANAA-ANQAVTTPAARALPLT 341	
RA	Garnier T., Eigmüller K., Camus J.-C., Medina N., Mansoor H.,		Db 293 GDAAKGLPGLGGMIG---GGPYAAGLONAASYGKLSSPPWNSGPLPSVTGAAPLPS 348	
RA	Pryor M., Durhoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,			
RA	Harris B., Atkin R., Doggett J., Keating L., Wheeler P.R.,		Qy 342 SLTAAERSPGQMLGGPVQMGARAGGGLSGYLVRPPRPPYMPHSAG 391	
RA	Parikh J., Barriel B.G., Cole S.T., Gordon S.V., Hewinson R.G.;		Db 349 TVSAAPEAAPSGLGLPJ---AGAGGAGAGP-RYGRFRPTMARPPFG 393	
RT	"The complete genome sequence of Mycobacterium bovis".;			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).			
DR	EMBL; BX248346; CAD95692.1; -.			
KW	Complete proteome.			
SQ	SEQUENCE 393 AA;	39413 MW;	RESULTS 8	
			05339 PRELIMINARY;	
O7TWF5		PRELIMINARY;	05339 PRELIMINARY;	
ID	Q7TWF5;		AC 05339 PRELIMINARY;	
AC			DT 01-JUN-1998 (TREMBLrel. 06, Created)	
DT	01-JUN-1998 (TREMBLrel. 06, Created)		DT 01-MAR-2003 (TREMBLrel. 24, Last sequence update)	
RA			DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
RA			DE PPE family protein.	
RA			GN RV1789 OR MT1838 OR MT1849.11.	
RA			OS Mycobacterium tuberculosis.	
RA			OC Bacteria; Actinobacteria; Actinomycetales; Mycobacterium.	
RA			NCBI_TaxID=1735;	
RA			RN SEQUENCE FROM N.A.	
RA			RP STRAIN=H37Rv;	
RA			RC MEDLINE=8295987; PubMed=9634230;	
RA			RA Cole S.T., Broscic R., Parkhill J., Garnier T., Churcher C., Harris D.,	

RESULT 7

Q7TZJ3 PRELIMINARY; PRT; 393 AA.

ID Q7TZJ3; AC Q7TZJ3; DT 01-OCT-2003 (TREMBLrel. 25, Created)



RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence";  
RL Nature 393:537-544 (1998).  
[2]

SEQUENCE FROM N.A.  
STRAIN=CDC 1551 / Oshkosh;  
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
Peterson J., DeBoy R., Dodson R., Gwinn M., Hafft D., Hickey E.,  
Kolonay J.F., Nelson W.C., Umayam J.A., Brmolaeva M., Salzberg S.L.,  
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mukul A.,  
Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains.,";  
Submitted to the EMBL/GenBank/DDBJ databases.  
EMBL: AL022021; CAA17729.1; ALT\_INIT.  
DR EAO07044; AAK46129.1; ALT\_INIT.  
PIR: A70932; A70932.  
TIGR: MT1856.1; -.  
DR Tuberculist; Rv1808; -.  
InterPro: IPK00003010; Microbac\_PBE.  
Pfam: PF00823; PPE; 1.  
Complete proteome.  
SEQUENCE 409 AA; 39917 MW; 1E15402BACF36379 CRC64;

Query Match 38.5%; Score 751; DB 16; Length 409;  
Best Local Similarity 44.7%; Pred. No. 8.4e-35;  
Matches 177; Conservative 53; Mismatches 138; Indels 28; Gaps 10;  
Qy 2 VDFGALPPEINSARMYAGPGSASILVAAQMDSVASDLSASAFQSVWGLTVGSWIGS 61  
Db 1 MDFGALPPEINSARMYAGPGSASILVAAQMDSVASDLSASAFQSVWGLTVGSWIGS 61  
Qy 62 SAGLMVAASAPYAWMSVTAQGELTAQVRVAAAYETAYGLTVPPVIAENRAELML 121  
Db 61 SSITMAAAYAPYAWMSVTAQGELTAQVRVAAAYETAYGLTVPPVIAENRAELML 120  
Qy 62 SAGLMVAASAPYAWMSVTAQGELTAQVRVAAAYETAYGLTVPPVIAENRAELML 121  
Db 61 SSITMAAAYAPYAWMSVTAQGELTAQVRVAAAYETAYGLTVPPVIAENRAELML 120  
Qy 122 IATNLIGONTPAIAVNEAEYGENWAQDAAMFGYAATATAATLIPFEEAPEMTSAGGL 181  
Db 121 VAINIEFGONTPAIAAATEHYAENWAQDAAMFGYAATATAATLIPFEEAPEMTSAGGL 179  
Qy 182 LEQAAAVEPASDTAA-----NOLMNVPQLQQLAQOPTQGTTPSKLGWKTIVSPH 234  
Db 180 AAQSAVVAQAGAAASSDTAQLSQLISLLPSLQLSIA-----NOLMNVPQLQQLAQOPTQGTTPSKLGWKTIVSPH 234  
Qy 235 RSPISNMYSMANTHMSMTNSGV-----MTNLSSMLKGFAFPAAAAQAVQTAQONGVRAM 289  
Db 233 LOSITTLANLTGPYSIGLGAIPGGMWMLFGQIL-GLAQNAFGVAAILGPKAAGALSP 292  
Qy 290 SSL-GSSLGS-SCLGGGVAANLGRAASYGSISVQWAAMANQAQVTPAARALPLTSLSA- 346  
Db 293 APLRGYGYIDITPLGGATGGIARATIVGSISVPOGWAEEAPVMRAVASVLPGTGAAPAL 352  
Qy 347 AERGPQMGJGLPVGQNGARAGGGL---SGVLRV 377  
Db 353 AADAPGALFGEMALLSLAGRALAGTAVRSGGAAARV 388

RESULT 11  
Q7TZH7 PRELIMINARY; PRT; 409 AA.  
AC Q7TZH7; PRELIMINARY; PRT; 399 AA.  
DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE PPE family protein.  
GN PPE1.0 OR MB1836.  
OS Mycobacterium bovis.  
OC Corynebacteriaceae; Actinomycetales;  
CC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1765;  
RN [1] \_  
RP SEQUENCE FROM N.A.

RC STRAIN=AF2122/97; PubMed=12788972;  
RX MEDLINE=22709107; Garnier T., Eiglemeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Grondin S., Lacroix C., Monsenpe C., Simon S.,  
RA Pryor M., Dutchoy S., Atkin R., Keating L., Wheeler P.R.,  
RA Harris B., Parkhill J., Barrell B.G., Cole S.T., Doggett J., Mayes R.,  
RA RT "The complete genome sequence of Mycobacterium bovis.,"  
Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
DR EMBL: BX248340; CAD94539.1; -.  
KW Complete proteome.  
SQ SEQUENCE 399 AA; 38840 MW; 1A0F43773187E74F2 CRC64;  
RN [1] \_  
RP SEQUENCE FROM N.A.

Query Match 38.4%; Score 749; DB 16; Length 409;  
Best Local Similarity 44.3%; Pred. No. 1.1e-34;  
Matches 176; Conservative 51; Mismatches 149; Indels 30; Gaps 10;  
Qy 2 VDFGALPPEINSARMYAGPGSASILVAAQMDSVASDLSASAFQSVWGLTVGSWIGS 61  
Db 1 MDFGALPPEINSARMYAGPGSASILVAAQMDSVASDLSASAFQSVWGLTVGSWIGS 61  
Qy 62 SAGLMVAASAPYAWMSVTAQGELTAQVRVAAAYETAYGLTVPPVIAENRAELML 121  
Db 61 SSITMAAAYAPYAWMSVTAQGELTAQVRVAAAYETAYGLTVPPVIAENRAELML 120  
Qy 122 IATNLIGONTPAIAVNEAEYGENWAQDAAMFGYAATATAATLIPFEEAPEMTSAGGL 181  
Db 121 VAINIEFGONTPAIAAATEHYAENWAQDAAMFGYAATATAATLIPFEEAPEMTSAGGL 179  
Qy 182 LEQAAAVEPASDTAA-----NOLMNVPQLQQLAQOPTQGTTPSKLGWKTIVSPH 234  
Db 180 AAQSAVVAQAGAAASSDTAQLSQLISLLPSLQLSIA-----NOLMNVPQLQQLAQOPTQGTTPSKLGWKTIVSPH 234  
Qy 235 RSPISNMYSMANTHMSMTNSGV-----MTNLSSMLKGFAFPAAAAQAVQTAQONGVRAM 289  
Db 233 LOSITTLANLTGPYSIGLGAIPGGMWMLFGQIL-GLAQNAFGVAAILGPKAAGALSP 292  
Qy 290 SSL-GSSLGS-SCLGGGVAANLGRAASYGSISVQWAAMANQAQVTPAARALPLTSLSA- 346  
Db 293 APLRGYGYIDITPLGGATGGIARATIVGSISVPOGWAEEAPVMRAVASVLPGTGAAPAL 352  
Qy 347 AERGPQMGJGLPVGQNGARAGGGL---SGVLRV 377  
Db 352 LAABAPGALFGEMALLSLAGRALAGTAVRSGGAAARV 388

RESULT 12  
Q7TZH8 PRELIMINARY; PRT; 399 AA.  
AC Q7TZH8; PRELIMINARY; PRT; 399 AA.  
DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE PPE family protein.  
GN PPE1.0 OR MB1836.  
OS Mycobacterium bovis.  
OC Corynebacteriaceae; Actinomycetales;  
CC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1765;  
RN [1] \_  
RP SEQUENCE FROM N.A.

RC STRAIN=AF2122/97; PubMed=12788972;  
RX MEDLINE=22709107; Garnier T., Eiglemeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Grondin S., Lacroix C., Monsenpe C., Simon S.,  
RA Pryor M., Dutchoy S., Atkin R., Keating L., Wheeler P.R.,  
RA Harris B., Parkhill J., Barrell B.G., Cole S.T., Doggett J., Mayes R.,  
RA RT "The complete genome sequence of Mycobacterium bovis.,"  
Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
DR EMBL: BX248340; CAD94539.1; -.  
KW Complete proteome.  
SQ SEQUENCE 399 AA; 38840 MW; 1A0F43773187E74F2 CRC64;  
RN [1] \_  
RP SEQUENCE FROM N.A.

Best Local Similarity 42.8%; Pred. No. 3.2e-34; Matches 182; Conservative 60; Mismatches 122; Indels 61; Gaps 13;

QY 2 VDFGALPPRENSARMYAGPGSASLVAACMWDSVASYLSSAASFOSVWGLTVGSWIGS 61  
 DB 1 MDFTALPPEINSARMYSGGSAPMLAASAWHGSIAERASALSYSSVLTSTGEHWGP 60

QY 62 SAGIMVAAASPYAAMSVTAAQVTAELAAQVRAAAAYETAYGLTPPPVIAENRAEMIL 121  
 DB 61 ASASNTAAASPYAAMSVTAAQVRAAAAYAAFAFATPPVIAENRAQINAL 120

QY 122 IATNLGQNTPAIAVNEAFYGENMADAAAMFGYAATAATATATLPPFEEADEMTSAGGL 181  
 DB 121 IATNLGQQAPATATEAQYAENWSODAMAMYAGASAAT QLTFTEPPVOTTNASGL 179

QY 182 LEQLAAVEEADSITAA----NOLMANYPQALQQLAOPTQGTT--PSSKUG-----225  
 DB 180 AAQSAAIATGAGAQQTTLSSOLIAAPSGLGIGLIGLGSGS 239

QY 226 ---GLWKTVSPHRSPISNNVMSMANNHMNTNSGVSM-TNLS---SMIKGFAPAAAQAV 278  
 DB 240 WLDKLWALDPN-----SNFWNTIASSGLFLPNTIAFPFLGLGGVAAADAQDV 289

QY 279 CTAAGONGVRAAMSISSL-----GSSGICGGVANLGRASVGSLSVPOQAWAAANOQAVP 333  
 DB 290 LGEATSG----GLGGALVAPLOSSAGGGTGTAGLGNATVGLTSVPSWTAAPIASP 344

QY 334 AARAL---PLTSLTSAAEKGPGQMGLGGLPGVONGARAGCGLGSVLRVPP---RPPYMPH 386  
 DB 345 LGSALGGTMVAPPBAAVAG---MPGMFGETNGGQFRG-----RAVPQYGRPNFVAR 394

QY 387 SPAG 391  
 DB 395 PPAG 399

RESULT 13  
 ID OS1956 PRELIMINARY; PRT; 403 AA.  
 AC 051956;  
 DT 01-JUN-1998 [TREMBUREL. 06, Created]  
 DT 01-JUN-1998 [TREMBUREL. 06, Last sequence update]  
 DE PBP-family protein  
 GN RV1807 OR MT1856 OR MTv049.29.  
 OC Mycobacterium tuberculosis.  
 Bacteria: Actinobacteria: Actinomycetidae: Actinomycetales:  
 NCBI\_TaxID=1773;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigleman K., Gas S., Barry C.E., III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Barrell B.G.;  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2] SEQUENCE FROM N.A.  
 RP STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Allard D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Hatt D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umeyama M., Brimble M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and

RT laboratory strains"; Submitted (APR 2001) to the EMBL/GenBank/DDBJ databases.  
 RL DR CAA7728.1;  
 DR DR EMBL AE007044; AAC46128.1;  
 DR PIR; H70931; H0931.  
 DR DR TIGR; MT1856;  
 DR DR InterPro; IPR00030; Microbac\_PPE.  
 KW DR Complete proteome.  
 FT DR CONFFLICT 227 227 F -> S (IN REF. 2).  
 FT DR CONFLICT 238 238 V -> L (IN REF. 2).  
 SQ DR CONFLICT 403 AA; 39243 MW; DCB18880FD15-BFE CRC64;

Query Match 37.8%; Score 737.5; DB 16; Length 403;  
 Best Local Similarity 42.6%; Pred. No. 4.8e-34;  
 Matches 181; Conservative 60; Mismatches 123; Indels 61; Gaps 12;

QY 2 VDFGALPPRENSARMYAGPGSASLVAQNMWDVSASDLESAASAFAQSYVWGLTVGSWIGS 61  
 DB 5 LDFTALPPEINSARMYSGGSSAPMLAASAWHGSIAERASALSYSSVSLSTLGEBWHGP 64

QY 62 SAGIMVAAASPYAAMSVTAAQVRAAAAYETAYGLTPPPVIAENRAELMIL 121  
 DB 65 ASASNTAAASPYAAMSVTAAQVRAAEQAGAAEAAAYAAFAATVPPVIEANPAQML 124

QY 122 IATNLGQNTPAIAVNEAFYGENMADAAAMFGYAATAATATLPPFEEADEMTSAGGL 181  
 DB 125 IATNLGQNAIAATEAQIADMVQGAGAAAT-QLTFTEPVQTINASGL 183

QY 182 LEQAAVEEADSITAA-----NOLMNNYPOQALQLAOPQGT-----TPSS 222  
 DB 184 AAQSAAIATGAGAQQTTLSSQIJAATPSVLQCLSSSTAATFASGPSLLGVGSQSS 243

QY 223 KLGGLKTVSPHRSPLSNKYSMANTHMSMTNSGISM-TNTLS---SMIKGFAPAAAQAV 278  
 DB 244 WLDKLWALDPN-----SNFWNTIASSGLFLPNTIAFPFLGLGGVAAADAQDV 293

QY 279 QTAAGONGVRAAMSISSL-----GSSGICGGVANLGRASVGSLSVPOQAWAAANQAVP 333  
 DB 294 LGERTG-----GLGSGALVAPLGSQGLGIGLTVAAIGNATVGLSVPSPSWTAAPIASP 348

QY 334 AARAL---PLTSLTSAAEKGPGQMGLGGLPGVONGARAGGGLSGVLRVPP---RPFYMPH 386  
 DB 349 LGSALGGTMVAPPVAAG---MPGMFCTMGQQGFG----RAVFQYGRPNFVAR 398

QY 387 SPAAG 391  
 DB 399 PPAG 403

RESULT 14  
 ID OS3950 PRELIMINARY; PRT; 423 AA.  
 AC 053950;  
 DT 01-JUN-1998 [TREMBUREL. 06, Created]  
 DT 01-JUN-1998 [TREMBUREL. 06, Last sequence update]  
 DT 01-JUN-2003 [TREMBUREL. 24, Last annotation update)  
 DE PPE-family protein.  
 GN RV1801 OR MTv049.23 OR MTv049.29.  
 OC Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCB\_ TaxID=1773;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=H37RV;  
 RC MEDLINE=92295987; PubMed=9634230;  
 RA Cole S.T., Broch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigleman K., Gas S., Barry C.E., III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Murphy L.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis clinical and

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,  
 RA Shilston J.E., Taylor K., Whitenhead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 [2]

RP SEQUENCE FROM N.A. STRAIN AR2122/97; PubMed=12788972;

RC STRAIN=CDC 1551 / Oshkosh; Garnier T., Biglmeier K., Camus H.-C., Medina N., Mansoor H.,

RA Fleischmann R.D., Allard D., Eisen J.A., Carpenter L., White O., Grondin S., Lacroix C., Simon S.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haff D., Hickey E., Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R., Parikh J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mkulula A., Parikh J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;

RA "The complete genome sequence of Mycobacterium bovis.";  
 DR Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
 DR EMBL; BX248310; CDD9452.1; -.

KW Complete proteome;  
 SQ SEQUENCE 423 AA; 41477 MW; 26E52CC271FBBF57 CRC64;

Query Match 37.8%; Score 736; DB 16; Length 423;  
 Best Local Similarity 41.8%; Pred. No. 6.2e-34; Indels 80; Gaps 12;  
 Matches 182; Conservative 50; Mismatches 123;

QY 2 VDFGALPPEINSARMYAGPGSASIVAAQMWDVSADLFSASAFOSVWGLTVGSWIGS 61

Db 1 MDFGALPPEINSARMYAGPGSASIVAAQMWDVSADLFSASAFOSVWGLTVGSWIGS 61

QY 62 SAGLMVAAASPYAWMSYTAGQELTAQVRYVAAYETAYGLTPPPVIAENRAELMIL 121

Db 60 SSTSMASAAASPYAWMSATAVHAELLAGQARIAAYAAATAFWAELAVEHATAGYASELSALT-GAVSGP 59

QY 122 IATNLGQNTPATAVNEAEYGENWQDAAAMEGYAAATATAATATATLPEEAPEMTSAGGL 181

Db 120 IATNIFGGQNTPATAMMEYQYMEMWQDAAAMTYGAGSSATA-SRMTAFTEPQQTINHQL 178

QY 182 LEQAAAVEBASDTAAAN-----QLMNIVPQALQIQLQPT----QGTTP-----S 221

Db 179 GAQSSAVAQATAAFAAGNLIQOSAFPQLLSAVPRALQGLPLTASQASATPQWVTDLGNS 238

QY 222 SKLGLWLWTKVSPHRSPISNMVSMANHNSMTNSGVSMINTLSSMLKGFAASVGSL 281

Db 239 TFLGG---AVTGPFYTFP-----GVLPSSGPVYPLGIQSVL-----V 271

QY 282 AONGVRAMSSLGS-----SLGSSGLGG-VANLGRAASVGSL 319

Db 272 TQNGGVSALLGKIGKPITGAIAPLAEPHTPILSEGGJGGSYSAGIRAGLYGKLS 331

QY 320 VPQAWAAANOQAVTPAARALPLTSLS--AAERGPQGMIGGLPVGQMGARAGGLGSGVLR 376

Db 332 VPQEWTVAAPEIPSPAAIQTARLAAAPTAATDGAGLGGMALSGLAAGSTG--- 388

QY 377 VPPRPYVNPBHSRAG 391

Db 389 ---HPIGSAAAAPAVG 400

Search completed: August 25, 2004, 03:13:02  
 Job time : 121 secs

RESULT 15  
 QTTZ14 PRELIMINARY;  
 ID QTTZ14  
 AC QTTZ14  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE PPE family protein.  
 GN PPE29 OR MB1829.

Database :		SwissProt - 42 : *				
Title:	US-09-724-685-107					
Run on:	August 25, 2004, 02:57:58 ; Search time 25 Seconds (without alignments)					
Scoring table:	BLOSUM62					
Sequence:	1 MVDFGALPBEINSARMYAGP..... SGVLRVPPRPTYMPHSPAAG 391					
Post-processing:	Minimum Match 0% Maximum Match 100% Listing First 45 summaries					
Searches:	141681 seqs, 52070155 residues					
Total number of hits satisfying chosen parameters:	141681					
Minimum DB seq length: 0	Maximum DB seq length: 2000000000					
Gapopen: 10.0 , Gapext: 0.5						
SUMMARIES						
Result No.	Score	Query Match	Length	DB ID	Description	
1	1655.5	85.0	396	1 YD61_MYCCTU	Q11031 mycobacter	
2	705	36.2	408	1 YS92_MYCCTU	Q10813 mycobacter	
3	688	35.3	463	1 YI02_MYCCTU	Q53951 mycobacter	
4	444	22.8	487	1 Y442_MYCCTU	P42611 mycobacter	
5	426.5	21.9	443	1 Y978_MYCCTU	Q10540 mycobacter	
6	418	21.4	678	1 YF48_MYCCTU	Q10778 mycobacter	
7	364.5	18.7	408	1 SRA_MYCULE	Q07297 mycobacter	
8	324.5	16.6	463	1 Y936_MYCCTU	Q10892 mycobacter	
9	324	16.6	434	1 YU18_MYCCTU	P31500 mycobacter	
10	321.5	16.5	435	1 YU21_MYCCTU	Q53268 mycobacter	
11	233.5	12.0	178	1 YY29_MYCCTU	Q06246 mycobacter	
12	217.5	11.2	176	1 YY25_MYCCTU	Q50703 mycobacter	
13	153.5	7.9	860	1 ELS_MOUSE	P54320 mus muscul	
14	151.5	7.8	232	1 YY26_MYCCTU	Q50702 mycobacter	
15	150	7.7	864	1 ELS_RAT	Q9372 ratto norvegicus	
16	143.5	7.4	730	1 ELS_HUMAN	P15502 homo sapiens	
17	139.5	7.2	881	1 PRYJ_YEAST	P47033 saccharomyces cerevisiae	
18	135	6.9	907	1 A1B0_HUMAN	Q8641 homo sapiens	
19	133.5	6.8	825	1 ICPO_HSV2H	P28284 herpes simplex virus type 1	
20	133.5	6.8	2432	1 Y43R_IRV6	P18305 chilo iridovirus	
21	132	6.8	1120	1 STFR_ECOLI	P76072 escherichia coli	
22	132	6.8	2090	1 N214_HUMAN	P35658 homo sapiens	
23	131	6.7	836	1 VG26_BPM5	Q05233 mycobacter	
24	131	6.7	1150	1 APM_PIG	P21021 sus scrofa	
25	129	6.6	1783	1 RAA3_CHLRE	Q94ec4 chlamydophila pneumoniae	
26	128	6.6	779	1 SRP_DROME	P52172 drosophila melanogaster	
27	127.5	6.5	790	1 ANP_NOTCO	P24656 notochelata	
28	126	6.5	354	1 YAUG_SCHPO	Q10169 schizosaccharomyces pombe	
29	125.5	6.4	1211	1 BUN2_DROME	Q24523 drosophila melanogaster	
30	125	6.4	577	1 CST2_HUMAN	P33240 homo sapiens	
31	124	6.4	2090	1 HFC1_MEASU	P51611 mesocricetus auratus	
32	123.5	6.3	677	1 VBL6_VENOM	Q50199 venus flytrap	
33	122	6.3	622	1 VBL7_VENOM	Q50197 venus flytrap	

## ALIGNMENTS

Searched: 141681 seqs, 52070155 residues  
 Total number of hits satisfying chosen parameters: 141681  
 Minimum DB seq length: 0 Maximum DB seq length: 0  
 Minimum Match 0% Maximum Match 100%  
 Listing first 45 summaries

**Database :** SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

Result No.	Score	Query Match	Length	DB	ID	Description
						%
1	1656.5	85.0	396	1	YD61_MYCTU	Q11031 mycobacteri
2	705	36.2	408	1	YS92_MYCTU	Q10813 mycobacteri
3	688.5	35.3	463	1	YI02_MYCTU	Q05391 mycobacteri
4	444	22.8	487	1	Y442_MYCTU	P42611 mycobacteri
5	426.5	21.9	443	1	Y878_MYCTU	Q10540 mycobacteri
6	418	21.4	678	1	YF48_MYCTU	Q10778 mycobacteri
7	364.5	18.7	408	1	SRA_T4C1E	Q07297 mycobacteri
8	324	16.6	463	1	Y096_MYCTU	Q10832 mycobacteri
9	324	16.6	434	1	YU18_MYCTU	P31500 mycobacteri
10	321.5	16.5	435	1	YU21_MYCTU	Q053268 mycobacteri
11	233.5	12.0	178	1	YY29_MYCTU	Q06246 mycobacteri
12	217.5	11.2	176	1	YY25_MYCTU	P07033 mycobacteri
13	153.5	7.9	860	1	EILS_MOUSE	P54320 mymusculi
14	151.5	7.8	232	1	YY26_MYCTU	Q50702 mycobacteri
15	150	7.7	864	1	EILS_PAT	Q93372 rattus norv
16	143.5	7.4	730	1	EILS_HUMAN	P155502 homo sapien
17	139.5	7.2	881	1	PRY3_YEAST	P47033 sacharomyces
18	135	6.9	907	1	A180_HUMAN	Q06411 homo sapien
19	133.5	6.8	825	1	ICPO_HSV2H	P28284 herpes simp
20	133.5	6.8	2432	1	Y43R_RV6	P18305 chilo iride
21	132	6.8	1120	1	STFR_ECOLI	P76072 escherichia
22	132	6.8	2090	1	N214_HUMAN	P356538 homo sapien
23	131	6.7	836	1	VG26_DPM15	Q05223 mycobacteri
24	131	6.7	1150	1	APMO_PIG	P12021 sus scrofa
25	129	6.6	1783	1	BAA3_CHLRE	Q9fec4 chlamydomon
26	128	6.6	779	1	SRP_DROME	P52172 drosophila
27	127.5	6.5	790	1	ANP_NCITCO	P4856 norotachia
28	126	6.5	354	1	YAUG_SCHPO	Q10169 schizosacch
29	125.5	6.4	1211	1	BUN2_DROME	P24523 drosophila
30	125	6.4	577	1	CST2_HUMAN	P33240 homo sapien
31	124	6.4	2090	1	HFC1_MEASU	P51611 mesocricetus
32	123.5	6.3	677	1	YI36_MYCTU	P05597 mycobacteri
33	123	6.3	1130	1	YD13_MYCTU	P21220 mycobacteri

Query	PF00823; PPE; 1. Heterological protein. Complete proteome.	Result 1
Best Local Similarity	85.0%	Score 1656 5; DB 1; Length 396;
Matches	338; Conservative 18; Mismatches 34; Indels 7; Gaps 3;	From N.A.
Qy	1 MVDGALPPEINSARMYAGPGSASLVAAKQMDSYASDLSAASFQSYVWGLTVGSMG 60	RSP SEQUENCE FROM N.A.
Ddb	1 MVDGALPPEINSARMYAGPGSASLVAAKQMDSYASDLSAASFQSYVWGLTVGSMG 60	SPECIES=M.bovis; STRAIN=AFF122/97;
Qy	61 SAGIMVAASPYAWMSVTAGQELTAQVRVAAAYETAYGTVPVVIAENRAELMI 120	MEDLINE=22703107; PubMed=121788972;
Ddb	61 SAGIMVAASPYAWMSVTAGQELTAQVRVAAAYETAYGTVPVVIAENRAELMI 120	Priyar M., Duboy S., Biglmeier K., Camus J.-C., Medina N., Mansoor H., Grondin S., Lacroix C., Monsempe C., Simon S., Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R., Parikh J., Barrell B.G., Cole S.V., Gordon S.V., Hewinson R.G.; "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; J. Bacteriol. 184:5479-5490 (2002).
Qy	121 LIATNLGONTPAIAVNEAEYGENWAQDAAMFGEYAAATATATLPPFEAAPMTSAGG 180	[3]
Ddb	121 LIATNLGONTPAIAVNEAEYGENWAQDAAMFGEYAAATATATLPPFEAAPMTSAGG 180	SEQUENCE FROM N.A.
Qy	181 LLQQAAVEEADSDTAANOLMNNYQALQQLAQOPTQGTPSKIGLWKTSPHRSPISN 240	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
Ddb	181 LLQQAAVEEADSDTAANOLMNNYQALQQLAQOPTQGTPSKIGLWKTSPHRSPISN 240	CC
Qy	241 MVSMNHHMSMTNSGVSMNTLSSMLKGPPAAAQQAVYOTAQNGVRAMSS ---LGSSL 296	CC
Ddb	241 IVSNLNNHYSMTNSGVSMASLHEMLKGFP-AAAQAVETAAQNGVQAMSSLGSQQLSSL 299	CC
Qy	297 GSSGLGGYAAANLGRAASVGSLSYSPQANAAQAVTAAQARPLTSLSAEEGPQYMLG 356	CC
Ddb	300 GSSGLGGYAAANLGRAASVGSLSYSPQANAAQAVTAAQARPLTSLSAEEGPQYMLG 359	CC
Qy	357 GLPVGQMGARA-G-GGLSGVLRVPRPTVMPHSPPAAG 391	DR TuberList: Rv2892C; -
Ddb	360 GLPGLQTNSSGGFGVSNALRIPRATVMPHPAAG 396	DR InterPro: IPR00030; Microbac_PPE.
/ RESULT 2		
Qy	Y892_MYCCTU STANDARD; PRT; 408 AA.	DR Pfam: PF00823; PPE; 1.
Ddb	AC QI0813; DT 01-OCT-1998 (Rel. 34, Created) DT 15-DEC-1998 (Rel. 37, Last sequence update) DE 2003 (Rel. 42, Last annotation update)	DR Hypothetical PPE-Family protein Rv2892C/Mt2959/Mb2916C.
Qy	DE DE Mycobacterium tuberculosis, and	DR Rv2892C OR Mt2959 OR MB2916C.
Ddb	DS DS Mycobacterium tuberculosis, and	DR
Qy	DC DC Bacteria; Actinobacteria; Actinomycetales;	DR
Ddb	NCBI_TAXID=1773, 1765;	DR
Qy	DN SEQUENCE FROM N.A.	DR
Ddb	SPCIES=M.tuberculosis; STRAIN=H37Rv;	DR
Qy	NR MEDLINE=98225987; PubMed=9534230;	DR
Ddb	RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutten S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence"; Nature 393:537-544 (1998).	DR
Qy	RN SEQUENCE FROM N.A.	DR
Ddb	SPCIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;	DR
Qy	2C SPECIES=M.tuberculosis; STRAIN=CD1551 / Oshkosh;	DR
Ddb	2C MEDLINE=22616494; Altland D., Eisen J.A., Carpenter L., White O., Fleischmann R.D., Alm R., Eisen R., Dodson R., Gwynn M., Haft D., Hickey E., Peterson J., DeBoy R., "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; J. Bacteriol. 184:5479-5490 (2002).	DR
Qy	2C VPRBPPYWPNSPAG 391	DR
Ddb	2C 394 YGFRLATMQRPFAG 408	DR



MEDLINE=22206494; PubMed=12218036;

RX Fleischmann R.D., Alland D., Bissen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M., Hatt D., Hickey E., Kolonay F.J., Nelson W.C., Umayam L.A., Emilia S.L., Delcher A., Utterback J., Weidman J., Khouri H., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; J. Bacteriol. 184:1479-1540(2002).

RL -!- SIMILARITY: Belongs to the mycobacterial PPE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR M15467; AAA88235.1; ALT\_INIT.

CC DR PIR: A1021932; CAA17339.1; -.

CC DR PIR: A006948; ARK44681.1; -.

CC DR PIR: C70830; C70830.

CC DR TIGR: MT0458; -.

CC DR Tuberculosis; RV0422C; -.

CC DR InterPro; IPR000030; Microbac\_PPE.

CC DR Pfam; PF00822; PPE\_1.

CC DR Pfam; PF01469; Pentapeptide\_2; 5.

CC DR PIR: P002989; Mycobac\_PPE.

CC DR PIR: P002989; Pentapeptide\_2; 5.

CC DR PIR: P002989; Pentaapep.

KW Hypothetical protein; Complete proteome.

FT CONFLICT 40 40 E > K (IN REF. 2).

FT CONFLICT 96 96 I -> T (IN REF. 1).

FT CONFLICT 211 211 G -> GNNNG (IN REF. 1).

SQ SEQUENCE 487 AA; 47247 MW; 47245B316C8CF CRC64;

CC Best Local Similarity 22.8%; Score 444; DB 1; Length 487; Matches 135; Conservative 48; Mismatches 180; Indels 58; Gaps 12;

CC QY 4 FGALPPEINSARMYAGPSASLYAAQMDVSADLFSAASAFAQSYYWGLTGVSNIGSSA 63

CC Db 6 FAWLPPEINSALMFAPGSSPLIAATAWGELABELIASIAGSIVSELISGAWLGPSSA 65

CC Query Match 64 GLAYAAASPVYAMSVTGAQELTTAQYVRAAAYETAGTYPPPYTAERMAELMILIA 123

CC Db 66 AAMNAVATQYLANISTAQAEEAAAQAMATAFEALAAIVQPAVVAANGLMOLLAA 125

CC QY 124 TNLLGQNTPAIANNEAEGEMTAQDAAAMPYAAATATATTLPPPEAPMTSAG--- 179

CC Db 126 TNMFGQNAPALMPVEAAEQMVALDVAAAMAGTHFDASAAVQLAPNQQV-IRNLIDIG 183

CC QY 180 -GLIBQAAVEPASDITAAANQMLNNVQPAQLQLAQPTQGTTPSSKUG-----GLW 228

CC Db 184 KNGQINLGFGNTGSIGNNNNIGSGNTGTGNIGSGNTGSGNLGNLDGNIGFG 243

CC Query Match 229 KTVSS-----PHRSPISSNNMSMANNHMTNTGSMTTSLSMIKGFPAAAAQQAVQ 279

CC Db 244 NTCGGNIGFRGFGTDHQMGFGNSGSGN-IGFGNSGTGNGLFNS----- 287

CC QY 280 TAACQNGYRAMSSLGSSGGLGGVVAANLGRAASYGSLSYPOAANAAQAVTPAARALP 339

CC Db 288 GSGNIGIANSINGIGTGT---INAGLG---SAGSLNT-SFWAGMQAAAGSAAGS 340

CC QY 340 LTSLTSAAEERGPQOM----LGGLPVQMGARAG-GGLGSQLR--VPPRFYVMPHSPAA 390

CC Db 341 EAALVSSAGYATGGMSTAALSSGILASALGTTGLOHGANVLNSLNTVFAAPASAPV 400

CC QY 391 G 391

CC Db 401 G 401

Y878\_MYCTU STANDARD; PRT; 443 AA.

ID Y878 MYCTU ID:Q10540; STRAIN=H37RV; AC Q10540; AC 01-OCT-1996 (Rel. 34, Created) DT 01-OCT-1996 (Rel. 34, Last sequence update) DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical PPE-family protein Rv0878c/MTO901. GN RV0878 OR MTO901 OR MTCY31.06C. OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales; Corynebacteriaceae; Mycobacterium; NCBI\_TaxID:1773; NCBI\_TaxID:1773;

RN SEQUENCE FROM N.A. [1]

RC STRAIN=H37RV; MEDLINE=98295987; PubMed=9634230; RX Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Bigdeli K., Gas S., Barry C.B. III, Tekala F., Badcock R., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLan J., Moile S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutters S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature 393: 537-544 (1998).

RN SEQUENCE FROM N.A. [2]

RP SEQUENCER FROM N.A. [2]

RC STRAIN=CDC 1551 / Oshkosh; MEDLINE=22206494 ; PubMed=12219036; RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M., Hatt D., Hickey E., RA Koleny J.F., Nelson W.C., Umayam L.A., Salzberg S.L., DeJcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains." J. Bacteriol. 184:15479-15490 (2002).

CC !- SUBCELLULAR LOCATION: Integral membrane protein (Potential). CC !- SIMILARITY: Belongs to the mycobacterial PPE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL/ Z73101; CAA97385.1; -.

CC DR EMBL/ AB006977; AAK45143.1; ALT\_INIT.

CC DR PIR: C70780; C70780.

CC DR TIGR: MR0901; -.

CC DR TuberList; Rv0878C; -.

CC DR InterPro; IP00030; Microbac\_PPE.

CC DR EMBL/ AB006977; AAK45143.1; ALT\_INIT.

CC DR PIR: C70780; C70780.

CC DR PIR: PR00823; PPE\_1.

CC KW Hypothetical protein; Transmembrane; Repeat; Complete proteome.

CC FT TRANSMEM 15 35 POTENTIAL.

CC FT TRANSMEM 38 58 POTENTIAL.

CC FT TRANSMEM 59 79 POTENTIAL.

CC FT TRANSMEM 181 201 POTENTIAL.

CC FT DOMAIN 64 73 ALA-RICH.

CC FT DOMAIN 81 115 ALA-RICH.

CC FT DOMAIN 231 270 4 X 10 AA APPROXIMATE REPEATS.

CC SQ SEQUENCE 443 AA; 45392 MW; C58BE660767552 CRC64;

Query Match 21.9% Score 426.5; DB 1; Length 443; Best Local Similarity 31.2%; Prcd. No. 4.5e-18; Matches 125; Conservative 56; Mismatches 132; Indels 87; Gaps 14;

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

CC EMBL; Z74020; CAA98335\_1;  
DR EMBL; AB007046; ARK45866\_1; ALT\_INIT.  
DR PIR; A70762; A70742.  
DR TIGR; MT1599; -  
DR Tuberculist; RV1548C; -  
DR InterPro; IPR000030; Microbac\_PPE.  
DR InterPro; IPR002989; Mycobac\_Pentapep.  
DR Pfam; PF01469; Pentapeptide\_2; 11.  
DR Pfam; PF00823; PPE\_1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 14 34 POTENTIAL.  
FT TRANSMEM 180 200 POTENTIAL.  
FT CONFLICT 258 258 D->G (IN REF. 2).  
SQ SEQUENCE 678 AA; 66736 MW; 209F1593D5233A2 CRC64;

Query Match 21.4 %; Score 418; DB 1; Length 678;  
Best Local Similarity 31.9 %; Pred. No. 2.2e-17;  
Matches 106; Conservative 48; Mismatches 140; Indels 38; Gaps 7;

QY 2 VDFGALPPEINSARMYAGPGSASLVIYAQQMWDVSASDLEPAAVQWGLTVGSWIGS 61  
DB 1 MNFSVLPPEINSALMFGAGCPGMIAASAWTGLQDGLSAAASASAVTSQLAGSWQGP 60

QY 62 SAGLMVAAASPYVAMSVTAQAFELTTAAQVRAAAAYETAYGLTPPPVIAENRAELML 121  
DB 61 ASAANTGVAASYARLTAAQAEOAQGQQAASFAEEAALAAATVPGASAANRGFLRSI 120

QY 122 IATNLIGONTPAIAVNEAYGEMWAQDAAMFGYAAATAATATLPLFEEAPEMTSAGG- 180  
DB 121 VASNLLGQNAPAIAAVEAVYEQMWAAVAAMLGTGEASAVALSSTPFTPSAATPTGG 180

QY 181 -----LLQQAAVEEASDTAAANQMLNNVPQALQLQOPTQGTTPSSKLGMLKTVS 232  
DB 181 AVIIAGFPFLDGNVTIGGNLGNLGNLGS ----FNPGSANTGSVNLN ----- 229

QY 233 PHRSPTSNM-VSMANNHMMTNSGYSMTNLSMKGFAAAPAAQAVQTAQNGTRAMS 290  
DB 230 -----ANTGPNLNSGNICSYNLLGNGNDLN-----PDSGNTGTLNWSGNIGSYN 276

QY 291 SIGSSIGSSGGGGYA-ANIGRAASVGSLSV 320  
DB 277 LGGGNISYNGNISGNTGDTNFG-GGNITGNLYN 307

RESULT 7

SRA\_MYCLE ID SRA MYCLE STANDARD; PRT; 408 AA.  
AC Q07237; DT 01-NOV-1995 (Rel. 32, Created)  
AC DT 01-NOV-1995 (Rel. 32, Last sequence update)  
AC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
AC DE Serie-rich antigen (25k) (45 kDa protein).  
AC SRA OR ML041 OR MLUCL383\_14.

OS Mycobacterium leprae.  
OC Corynebacteriaceae; Mycobacterium.  
OC Corynebacterineae; Mycobacterium.  
NCBI\_TaxID=1769;

SEQUENCE FROM N.A.

STRAIN=H37RV; SEQUENCE=9829987; PubMed=9634230; RXN 291 SIGSSIGSSGGGGYA-ANIGRAASVGSLSV 320  
MEDLINE=22206494; PubMed=12218036; Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., RA Gordon S.V., Eiglmeier K., Barry C.E. III, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B. G.; RA complete genome sequence."; RT complete genome sequence from the Nature 39:537-544(1998).  
RN [1] RT [2]

SEQUENCE FROM N.A.

STRAIN=CDC 1551 / Oshkosh; RXN 291 SIGSSIGSSGGGGYA-ANIGRAASVGSLSV 320  
MEDLINE=22206494; PubMed=12218036; Carpenter L., White O., Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwynn D., Hickey E., RA Kolonay J.F., Nelson W.C., Uniyal L.A., Ermolaeva M., Salzberg S.L., Delcher A., Utterback R., Weidman J., Khouri H., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; RT J. Bacteriol. 184:5479-5490(2002).  
RL - SIMILARITY: Belongs to the mycobacterial PPE family.  
CC [2]

SEQUENCE FROM N.A.

STRAIN=CDC 1551 / Oshkosh; RXN 291 SIGSSIGSSGGGGYA-ANIGRAASVGSLSV 320  
MEDLINE=22206494; PubMed=12218036; Carpenter L., White O., Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwynn D., Hickey E., RA Kolonay J.F., Nelson W.C., Uniyal L.A., Ermolaeva M., Salzberg S.L., Delcher A., Utterback R., Weidman J., Khouri H., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; RT J. Bacteriol. 184:5479-5490(2002).  
RL - SIMILARITY: Belongs to the mycobacterial PPE family.  
CC [2]

SEQUENCE FROM N.A.

STRAIN=CDC 1551 / Oshkosh; RXN 291 SIGSSIGSSGGGGYA-ANIGRAASVGSLSV 320  
MEDLINE=22206494; PubMed=12218036; Carpenter L., White O., Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwynn D., Hickey E., RA Kolonay J.F., Nelson W.C., Uniyal L.A., Ermolaeva M., Salzberg S.L., Delcher A., Utterback R., Weidman J., Khouri H., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; RT J. Bacteriol. 184:5479-5490(2002).  
RL - SIMILARITY: Belongs to the mycobacterial PPE family.  
CC [2]

SEQUENCE FROM N.A.

STRAIN=CDC 1551 / Oshkosh; RXN 291 SIGSSIGSSGGGGYA-ANIGRAASVGSLSV 320  
MEDLINE=22206494; PubMed=12218036; Carpenter L., White O., Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwynn D., Hickey E., RA Kolonay J.F., Nelson W.C., Uniyal L.A., Ermolaeva M., Salzberg S.L., Delcher A., Utterback R., Weidman J., Khouri H., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; RT J. Bacteriol. 184:5479-5490(2002).  
RL - SIMILARITY: Belongs to the mycobacterial PPE family.  
CC [2]

RX MEDLINE=90202054; PubMed=7934645;	OY 214 ---PTQ-----GTPSSKLGLWKTIVSPHRSPIANNVSMANHNSMTNSGVMTNT 261
RA Rinke de Wit T.P., Clark-Curtiss J.E., Abebe F., Kolkon A.H.J.,	DR 240 PFGTPSSQSNDLSATSLTQQLGSL-----DSIASSLTLNTS--ISSST 286
RA Jonson A.A.M., Thole J.E.R.,	DR 262 LSSMLKGAFAPAAAQVQTAQGLGSGVAAAGGGVAAAN----LGRAASVG 316
RT "A Mycobacterium leprae-specific gene encoding an immunologically	DR 287 ASSIM---PIVASQVTEIGRSQV-AVEKMIQSISSTAVSVDAASSVTVAGVQARGGGLSGV 341
RT recognized 45 kDa Protein."	DR 317 SLSVPQAWAANAOAVTAAPALP-LTSLTSAERGPQMLGGHPVGQMGARAGGGLSGV 374
RL Mol. Microbiol. 10: 829-338 (1993).	DR 342 ALRVPEWATASQPVMATAHSVPGCSAITA-----VSGPLSGV 381
RN [3]	
RP SEQUENCE FROM N.A.	
RC STRAIN=TN;	
RX MEDLINE=1128732; PubMed=11234002;	OY 3175 LRVPFRPYMPHSPAG 391
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,	DR 382 TQ--PAEVLTAASVAGG 396
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,	
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,	
RA Davies R.M., Devlin K., Dutchoy S., Fraser A., Hamlin N.,	
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., MacLean J., Moule S.,	
RA Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,	
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,	
RA Squares S., Stevens K., Taylor K., Whittlehead S., Woodward J.R.,	
RA Barrell B.G.,	
RT "Massive gene decay in the leprosy bacillus.";	
RL Nature 409:1007-1011(2001).	
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC or send an email to license@isb-sib.ch).	
CC	
DR U0015; AAC3220_1; -.	OY 9096 MYCTU STANDARD; PRT; 463 AA.
DR X6A431; CAA8480_1; -.	AC Q10832; PRT; 463 AA.
DR EMBL; 221952; CAA79950_1; -.	AC Q10832; PRT; 463 AA.
DR EMBL; 297179; CAB79938_1; -.	DR 01-OCT-1996 (Rel. 34, Last sequence update)
DR EMBL; AL583948; CAC29919_1; -.	DR 10-OCT-2003 (Rel. 42, Last annotation update)
DR PIR; C86960; C86960.	RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
DR PIR; S33522; S33522.	RA Davies R., Devin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
DR PIR; S39872; S39872.	RA Gordon S.V., Eiglmeier K., Gas S., Tekala E.,
DR Lepronra; ML0411; -.	RA Hornsby T., Jagels K., Krogh A., McLean J., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
DR InterPro; IPR000030; Microbac_PPE.	RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
DR PF0223; PPE; 1.	RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
XR Antigen; Repeat; Complete proteome.	RA "Deciphering the biology of Mycobacterium tuberculosis from the
FT DOMAIN 192 196 POLY-SER.	RT complete genome sequence.";
FT DOMAIN 209 235 2 X 6 AA REPEATS OF S-V-A-Q-S-E.	RT RN 393:537-544 (1998).
FT REPEAT 209 235 2 X 6 AA REPEATS OF S-V-A-Q-S-E.	RT RN 393:537-544 (1998).
FT REPEAT 209 235 2 X 6 AA REPEATS OF S-V-A-Q-S-E.	RT RN 393:537-544 (1998).
FT REPEAT 209 235 2 X 6 AA REPEATS OF S-V-A-Q-S-E.	RT RN 393:537-544 (1998).
FT CONFLICT 132 132 T > S (IN REF. 2).	RT RN 393:537-544 (1998).
FT CONFLICT 189 189 S > R (IN REF. 2).	RT RN 393:537-544 (1998).
FT CONFLICT 191 191 H > D (IN REF. 2).	RT RN 393:537-544 (1998).
FT CONFLICT 292 292 P > L (IN REF. 2).	RT RN 393:537-544 (1998).
SQ SEQUENCE 408 AA; 42466 MW; 5C0C2BE0D6A98 CRC64;	RT RN 393:537-544 (1998).
Query Match 18.7% Score 364.5; DB 1;	
Best Local Similarity 26.3%; Pred. No. 1.7e-4;	
Matches 115; Conservative 72; Mismatches 163; Indels 87; Gaps 11;	
QY 1 MDFGALPPEINSARMYAGPGSASLYAAQWQWDYASDLFASASAQFQSVVWGLTVGSWIG 60	
Db 1 MDFGALPPEINSARMYAGPGSASLYAAQWQWDYASDLFASASAQFQSVVWGLTVGSWIG 60	
Qy 61 SSAGLMWAAASPVAWSVTAAQVRAAAYETAYGLTVPPVTAENFAELMI 120	
Db 61 EISDMILASRSVTFVAWLGNAAENAGLIARYLVHAYAEEFRAGMPLTGLNLIHTMA 120	
Qy 121 LIATNLQGQNTPAIAYNEAYGEMWAQDAAMFGIAAATATAATLPPREAPENTSAGG 180	
Db 121 LIAINWQGVSTVVAEADLMMQNSTMATITRDVTRGRMFNEDAPQVSR-Y 179	
Qy 181 LLEQQAAVEEAESDTAAANQMNKNVPAQQLQAQ-----213	
Db 180 CMRRDSVNSFHSSSSDLSLIESIDNLYSVQAQBHGDSMSQSNTCGSSVAQSELCD5 239	



	298	GGPLIGALAAAAPVGAVAGLAGVAGLAAL-PAVGAA- -AGAPAAVLGVSVAPVSGGVSPQA	354
Db	324	WAANQATPPAARMLPLTSLTAERGQCMUQGLGPVQGMGRAGGGLGSV	374
Qy		:	:
Db	355	RLVS- -AVEPAPASTSVSVL- -ASDRGAGAL- GF- VGTAGKESVGOPAGL	398
	RESULT 1.0		
	YU21_MYCUTU	STANDARD;	PRT;
	ID _YU21_MYCUTU		435 AA.
	AC 053268 ;		
	AC 053268 ;	Rel. 40, Created)	
	DT 16-OCT-2001	(Rel. 40, Last sequence update)	
	DT 16-OCT-2001	(Rel. 40, Last annotation update)	
	DT 10-OCT-2003	(Rel. 42, Last annotation update)	
	HYPOTHEtical PPE_FAMILY		
	DE RV3021C/RV3022C	RV3021C/RV3022C/MT3106.	
	DE RV3021C/RV3022C	MT3106 OR MT3106 OR MT3106 OR MT3106.	
	LOC Mycobacterium tuberculosis.		
	ORG Bacteria; Actinobacteridae; Actinomycetales;		
	NCBI_TaxID=1773;		
	OX 111RN		
	RRP SEQUENCE FROM N.A.		
	RXC STRAIN H37Rv		
	RXC MEDLINE 982951987; PubMed-9634210;		
	RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,		
	RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tebaria F.,		
	RA Badcock K., Basham D., Brown D., Chillingworth T., Connor P.,		
	RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,		
	RA Hornsby T., Osborne J., Krough A., McLean J., Moule S., Murphy L.,		
	RA Oliver S., Rutter S., Seeger K., Skelton S., Squares S., Rogers J.,		
	RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;		
	RT "Deciphering the biology of Mycobacterium tuberculosis from the		
	RT complete genome sequence.";		

[2] NATURE 394:53 - 544 (1998).

SEQUENCE FROM N.A.  
 STRAIN=CDC 1551 / Oshkosh;  
 MEDLINE=22206194 ; PubMedID:12218036;  
 File schmann R.D., Allard D., Eisen J.A., Carpenter L., White O.,  
 Peterson J.Y., DeBoy R., Dodson R., Gwin M., Haft D., Hickie E.,  
 Kolonay J.F., Nelson W.C., Umayam L.A., Brinkac M., Salzberg S.L.,  
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M./  
 "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 laboratory strains";  
 J. BACTERIOL. 184:15479-15490 (2002).  
 -|- SIMILARITY: Belongs to the mycobacterial PPE family.  
 -|- CAUTION: Ref.1 sequence differs from that shown due to a  
 frameshift in position 82.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

---

EMBL; AL021287; CAA16106;1; ALT FRAME.  
 EMBL; AL021287; CAA16107;1; ALT FRAME.  
 EMBL; AE007129; AAK47435;1; -.  
 TIGR; MT3106; -.  
 Tuberculist; RV3021C; -.  
 Tuberculist; RV3022C; -.  
 InterPro; IPR000030; Microbac\_PBE.  
 Pfam; PF000823; PPE; 1.  
 Hypothetical protein; Complete proteome.  
 CONFLICT 299 G -> A (IN REF. 2).  
 CONFLICT 317 320 LAGV (IN REF. 2).  
 CONFLICT 326 326 I -> V (IN REF. 2).  
 SEQUENCE 435 AA; 42876 MW; 3B157643KEA8484A CRC64;

RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 laboratorial strains.";  
 RL J. Bacteriol. 184:15479-15490 (2002).

-!- SIMILARITY: Belongs to the mycobacterial PPE family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 295389; CAB007158; AAK47873.1; -.  
 DR PTR; C73975; C70955.  
 DR TuberCultist; Rv34249; -.  
 DR InterPro; IPR000050; Microbac\_PPE.  
 DR Pfam; PF00823; PPB; 1.  
 DR Hypothetical protein; Complete proteome.

KW Hypothetical protein; Complete proteome.

SEQUENCE 178 AA; 1981 MW;

8BE1FC025ABFBAA6 CRC64;  
 DR Query Match, Score 12.0%; Best Local Similarity 35.9%; Pred. No. 2. Be-07;  
 DR Matches 65; Conservative 25; Mismatches 76; Indels 15; Gaps 3;

DR Qy 7 LPPPEINSAMYAGGSSASISVAAAQ---MWDS---VADSLFGRASAFOQVWWGLTVGSWI 59  
 DR 5 IPAYIYSNIIYEGPAGDSLSAEEQLRIMNSANMTAKSLTRLGELQE-----NWK 56  
 DR Qy 60 GSSAGLMLVAASAPVAVMSVTAGGAELTAQVRVAALEYETAVGLTPPVIAENRAELM 119  
 DR 57 GSSSDLMIAAARGYLDWLTKHSQILEYKIVDYLAVYYEETHKVYPPATANNREBVH 116  
 DR Qy 120 ILIATNLGCTNTPAIAVNAEYGENMWAAQDAAMFGYAAATAATATLIPFEAEPEMTSAG 179  
 DR 117 RLIASNVAGNTPAITAGLDAQYQQYRAQNIANMDYQSTARFILAYLPWQEPPOIYGGG 176  
 DR Qy 180 G 180  
 DR Db 177 G 177

RESULT 12  
 YY25 MYCTU STANDARD; PRT; 176 AA.  
 ID YY25 MYCTU STANDARD; PRT; 176 AA.  
 AC Q50703; DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-DEC-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical PPE-family protein Rv3425.  
 GN RV3425 OR MTcY78\_04C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TaxID=1773;  
 RN [1]  
 RP STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Broich R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekia F.,  
 RA Badcock K., Besham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feitwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jigels K., Krogh A., McLean J., Moulé S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutten S., Seeger K., Skelton S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence.";  
 RT Nature 393:537-544 (1998).

-!- SIMILARITY: Belongs to the mycobacterial PPE family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 277165; CAB01031.1; -.  
 DR PIR; F70738; F70738.  
 DR TuberCultist; Rv3425; -.  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPB; 1.  
 DR Hypothetical protein; Complete proteome.

KW Hypothetical protein; Complete proteome.

SEQUENCE 176 AA; 1985 MW;

BCECF25A463BB7BD CRC64;

DR Query Match, Score 217.5%; Best Local Similarity 33.5%; Pred. No. 2.3e-06;  
 DR Matches 58; Conservative 31; Mismatches 83; Indels 1; Gaps 1;

DR Qy 7 LPPETNSARMYAGGSSASISVAAAQWDSVASYASAFQSYVWGTLTYGSWIGSSAGLM 66  
 DR 5 IPAYIYSNIIYEGPAGDSLSFAGSORELAYSVETAESDEELBD-ENWKQSSSDLL 63

DR Db 67 VAASSPYVAMSUTAGQAELTAQAVRVAAYAETAYGLTYPPTVTAENRAELMILIAATNL 126  
 DR 64 ADAVEYXLOWNSKHSQSKHSSQLEKAWVNGLNAYANDTRKVYPPEEAANRERRIASNV 123  
 DR Qy 127 LGQNTPAIANVNAEYCEMWAQDAAMFGYAAATAATATLIPFEAEPEMTSAG 17.9  
 DR 124 AGVNTPAIADLDAQYDQYRARNVATMNAVSWTRSAISDLPRWRREPQTYRGG 176  
 DR Db .

RESULT 13

ELS\_MOUSE STANDARD; PRT; 860 AA.

AC P54320; DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE Elastin precursor (Troponoelastin).

GN ELN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBITaxonID=10090;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Lung;

RX MEDLINE=7829060;

RA Wynder K.S., Sechler J.L., Boyd C.D., Passmore H.C.,

RT "Use of an intron polymorphism to localize the tropoelastin gene to

RT mouse chromosome 5 in a region of linkage conservation with human

RT chromosome 7.";  
 Genomics 23:125-131 (1994).

CC -!- FUNCTION: Major structural protein of tissues such as aorta and

CC mucosal ligament, which must expand rapidly and recover completely.

CC -!- SUBUNIT: The polymeric elastin chains are cross-linked together

CC into an extensible 3D network.

CC -!- SUBCELLULAR LOCATION: Extracellular matrix of elastic fibers.

CC -!- PTM: The crosslinks are made of deaminated Lys.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR U088210; AAR80155.1; -.





This Page Blank (Continued)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 25, 2004, 03:13:09 ; Search time 127 Seconds

(without alignments)

967.510 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDFGALPPEINSARMYAGP.....SGYLRVPERPYVMPHSPAAAG 391

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters:

1295152

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database :

```

Published Applications AA:*
 1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*
 2: /cgn2_6/ptodata/1/pubpa/PCN_NEW_PUB.pep:*
 3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep:*
 4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep:*
 5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep:*
 6: /cgn2_6/ptodata/1/pubpa/PTETUS_PUBCOMB.pep:*
 7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep:*
 8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep:*
 9: /cgn2_6/ptodata/1/pubpa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1949	100.0	391	12 US-09-886-349A-14	Sequence 14, App
2	1949	100.0	391	14 US-10-198-002-102	Sequence 102, App
3	1949	100.0	391	14 US-10-084-843-107	Sequence 107, App
4	1949	100.0	391	14 US-10-098-732A-14	Sequence 14, App
5	1949	100.0	596	9 US-09-287-849-56	Sequence 26, App
6	1949	100.0	596	12 US-03-886-349A-20	Sequence 20, App
7	1949	100.0	596	14 US-10-359-460-26	Sequence 26, App
8	1949	100.0	596	14 US-10-098-732A-20	Sequence 20, App
9	1949	100.0	600	9 US-09-287-849-22	Sequence 22, App
10	1949	100.0	600	14 US-10-359-460-22	Sequence 22, App
11	1949	100.0	723	15 US-10-369-983-2	Sequence 2, App
12	1949	100.0	729	14 US-03-886-349A-18	Sequence 18, App
13	1949	100.0	729	14 US-10-098-732A-18	Sequence 18, App
14	1949	100.0	729	15 US-10-369-983-21	Sequence 21, App
15	1949	100.0	729	15 US-10-369-983-22	Sequence 22, App

ALIGNMENTS

RESULT 1

US-09-886-349A-14

; Application US/09886349A

; Publication No. US20040086553A1

; GENERAL INFORMATION:

; APPLICANT: Skeky, Yasir

; REEDER, Steven

; ALDERSON, Mark

; CORTIX CORPORATION

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis

; FILE REFERENCE: Q14058-00970US

; CURRENT APPLICATION NUMBER: US/09-8866, 349A

; CURRENT FILING DATE: 2001-06-20

; PRIOR APPLICATION NUMBER: US/09-597-796

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: US/60/265, 737

; PRIOR FILING DATE: 2001-02-01

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO: 14

; LENGTH: 391

; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis

; FEATURE: OTHER INFORMATION: MTB39 (TBH9FL)

; US-09-886-349A-14

Query Match 100.0% Score 1949; DB 12; Length 391;

Best Local Similarity 100.0% Pred. No. 1.2e-145;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDGFADPPEINSARMYAGPGSASLVAQMDSVASDLSAAASATOSVNGILTVGSWIG 60

Db 1 MYDFGALPPEINSARMYAGPGSASLVAQMDSVASDLSAAASATOSVNGILTVGSWIG 60

Qy 61 SSAGLMYAAASPVAVNSVTGQAEITTAQVRVAAAYETAGLTVPPVIAENRAELMI 120

Db 61 SSAGLMYAAASPVAVNSVTGQAEITTAQVRVAAAYETAGLTVPPVIAENRAELMI 120

Query Match 100.0%; Score 1949; DB 14; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-145;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPEINSARNYAGPGSASLIVAAQMDVSDAISDIFSAASAFQSTVNGITVGSWIG 60  
 Db 1 MVDFGALPPEINSARNYAGPGSASLIVAAQMDVSDAISDIFSAASAFQSVWGLTVGSWIG 60

Qy 61 SSAGIMVAAAASPYYAVMSVTAQGAELTAQYRVAAAAYETAYGTYPVPIAVNRAELMI 120  
 Db 61 SSAGIMVAAAASPYYAVMSVTAQGAELTAQYRVAAAAYETAYGTYPVPIAVNRAELMI 120

Qy 121 LIATNLIGONTPAIAVNEARYGEMQAQDAAMFGEYAAATATAATATLPEEAEPMTSGG 180  
 Db 121 LIATNLIGONTPAIAVNEARYGEMQAQDAAMFGEYAAATATAATATLPEEAEPMTSGG 180

Qy 181 LLEQAAVEASDTAAANQMLNNPQALQLAQTPGTTPSKLGGMKTVSPHSPISN 240  
 Db 181 LLEQAAVEASDTAAANQMLNNPQALQLAQTPGTTPSKLGGMKTVSPHSPISN 240

Qy 241 MVSHDANNMSMTNSMVTSSMNLKGFPAAAAQVTAQNGVRAMSLSLOSSSG 300  
 Db 241 MVSHDANNMSMTNSMVTSSMNLKGFPAAAAQVTAQNGVRAMSLSLOSSSG 300

Qy 301 LGGGVAAANLGRAASYGSLSVPOAAVAAANOAVTAAPARALPLTSAAERGPQOMLGGIPV 360  
 Db 301 LGGGVAAANLGRAASYGSLSVPOAAVAAANOAVTAAPARALPLTSAAERGPQOMLGGIPV 360

Qy 361 QMGARAGGCLSGVLRVPPRPPYMPHSAAAG 391  
 Db 361 QMGARAGGCLSGVLRVPPRPPYMPHSAAAG 391

Qy 241 MVSMANNHEMMNTNSGYSMTNTSSMLKGFPAAAAQVYOTAQNGVRAMSLGSSG 300  
 Db 241 MVSMANNHEMMNTNSGYSMTNTSSMLKGFPAAAAQVYOTAQNGVRAMSLGSSG 300

Qy 301 LGGGVAAANLGRAASYGSLSVPOAAVAAANOAVTAAPARALPLTSAAERGPQOMLGGIPV 360  
 Db 301 LGGGVAAANLGRAASYGSLSVPOAAVAAANOAVTAAPARALPLTSAAERGPQOMLGGIPV 360

Qy 361 QMGARAGGCLSGVLRVPPRPPYMPHSAAAG 391  
 Db 361 QMGARAGGCLSGVLRVPPRPPYMPHSAAAG 391

RESULT 3  
 US-10-084-843-107  
 Sequence 107, Application US/1008443  
 Publication No. US2003014243A1  
 GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 Skeiky, Yasir A.W.  
 Dillon, Davin C.  
 Campos-Neto, Antonia  
 Houghton, Raymond  
 Veivick, Thomas S.  
 Twardzik, Daniel R.  
 Lode, Michael J.  
 Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
 NUMBER OF SEQUENCES: 350  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SPED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/193,002  
 FILING DATE: 10-Jul-2002  
 CLASSIFICATION: <Unknown>  
 NAME: Maki, David J.

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,596  
 FILING DATE: 05-MAY-1998  
 ATTORNEY/AGENT INFORMATION:  
 REFERENCE/DOCKET NUMBER: 31\_392  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-5031  
 INFORMATION FOR SEQ ID NO: 102:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 391 amino acids  
 TYPE: amino acid  
 STRANDDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-10-084-843-107  
 NUMBER OF SEQUENCES: 355  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SPED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/084,843  
 FILING DATE: 25-Feb-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,967  
 FILING DATE: 05-MAY-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31\_392  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-5031  
 INFORMATION FOR SEQ ID NO: 102:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 391 amino acids  
 TYPE: amino acid  
 STRANDDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 102:

REFERENCE/DOCKET NUMBER: 210121.411C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-1900  
 TELEFAX: (206) 681-6031  
 INFORMATION FOR SEQ ID NO: 107:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 30 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 107:  
 US-10-084-841-107

Query Match 100.0%; Score 1949; DB 14; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-145; Indels 0; Gaps 0;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLVAALQMWDSDVSAASAFQSYYWGLTVGSWIG 60  
 Db 1 MVDFGALPPEINSARMYAGPGSASLVAALQMWDSDVSAASAFQSYYWGLTVGSWIG 60

Qy 61 SSAGLMVAASPTAAMSUTAGQELTAQVRYAAAYTAYGTTTTPPVIAENRAELMI 120  
 Db 61 SSAGLMVAASPTAAMSUTAGQELTAQVRYAAAYTAYGTTTTPPVIAENRAELMI 120

Qy 121 LIATNLIGQNTPAAVNEAEYGENWAQDAAMFGYAATATAATLPLPEEAPMTSAGC 180  
 Db 121 LIATNLIGQNTPAAVNEAEYGENWAQDAAMFGYAATATAATLPLPEEAPMTSAGC 180

Qy 181 LLEQAAAVEASDTAAANQLMNVQALQQLAQPTGTTPSKLGLWKTVPSPRSPISN 240  
 Db 181 LLEQAAAVEASDTAAANQLMNVQALQQLAQPTGTTPSKLGLWKTVPSPRSPISN 240

Qy 241 MVSMANNHMSMTNSGVSMINTLSSMLKGFPAAAQQVOTAONGVRAMSSLGSSG 300  
 Db 241 MVSMANNHMSMTNSGVSMINTLSSMLKGFPAAAQQVOTAONGVRAMSSLGSSG 300

Qy 301 LGGGVAAANLGRAASVGSLSVPQAWAAANQAVTAERGPQMLGGLPV 360  
 Db 301 LGGGVAAANLGRAASVGSLSVPQAWAAANQAVTAERGPQMLGGLPV 360

Qy 361 GQM GARAGGGSGVLRVPPRPPYMPHSPIAG 391  
 Db 361 GQM GARAGGGSGVLRVPPRPPYMPHSPIAG 391

RESULT 5  
 US-09-287-849-26  
 Sequence 26, Application US/09287849  
 Patent No. US20020009459A1  
 GENERAL INFORMATION:  
 / APPLICANT: Reed, Steven G.  
 / INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 / TITLE OF INVENTION: and Their Uses  
 / FILE REFERENCE: 014058-00020US  
 / CURRENT APPLICATION NUMBER: US/09/287-849  
 / CURRENT FILING DATE: 1999-04-07  
 / PRIOR APPLICATION NUMBER: US 08/818,112  
 / PRIOR FILING DATE: 1997-03-13  
 / PRIOR APPLICATION NUMBER: US 08/942,578  
 / PRIOR FILING DATE: 1997-10-01  
 / PRIOR APPLICATION NUMBER: US 09/025,197  
 / PRIOR FILING DATE: 1998-02-18  
 / PRIOR APPLICATION NUMBER: US 09/056,556  
 / PRIOR FILING DATE: 1998-04-07  
 / PRIOR APPLICATION NUMBER: US 09/223,040  
 / PRIOR FILING DATE: 1998-12-30  
 / NUMBER OF SEQ ID NOS: 46  
 / SOFTWARE: Patentin Ver. 2.1  
 / SEQ ID NO: 26  
 / LENGTH: 596  
 / OTHER INFORMATION: Description of Artificial Sequence.bi-fusion  
 / US-09-287-849-26

Query Match 100.0%; Score 1949; DB 9; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 2e-145;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLVAALQMWDSDVSAASAFQSYYWGLTVGSWIG 60



Qy 241 MVSMMANNHSMNTNSGYSMTNTLSSMLKGFPAPAAAQQAVCTAAQNGVRAMSIGSSLGSSG 300  
 Db 249 MVSMMANNHSMNTNSGYSMTNTLSSMLKGFPAPAAAQQAVCTAAQNGVRAMSIGSSLGSSG 308

Qy 301 LGGGVANILGRAASVSSLVQAWAANQAVTAAQVTPARALPLTSAAEPRGPQMLGGLPV 360  
 Db 309 LGGGVANILGRAASVSSLVQAWAANQAVTAAQVTPARALPLTSAAEPRGPQMLGGLPV 368

Qy 361 GONGARAGGGSGLVLRVPPRYPVMPHSPAG 391  
 Db 369 GONGARAGGGSGLVLRVPPRYPVMPHSPAG 399

RESULT 8  
 US-10-098-732A-20  
 / Sequence 20, Application US/10098732A  
 / Publication No. US20030175294A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Brannon, Mark  
 / APPLICANT: Sheiky, Yasir A. W.  
 / APPLICANT: Dillon, Davin C.  
 / APPLICANT: Alderson, Mark  
 / APPLICANT: Campos Neto, Antonio  
 / APPLICANT: Corixa Corporation  
 / TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 / FILE REFERENCE: 014058-00920US  
 / CURRENT APPLICATION NUMBER: US/09/287-849  
 / PRIOR APPLICATION NUMBER: US 08/818,112  
 / PRIOR FILING DATE: 1997-03-13  
 / PRIOR APPLICATION NUMBER: US 08/942,578  
 / PRIOR FILING DATE: 1997-10-01  
 / PRIOR APPLICATION NUMBER: US 09/025,197  
 / PRIOR FILING DATE: 1998-02-18  
 / PRIOR APPLICATION NUMBER: US 09/056,556  
 / PRIOR FILING DATE: 1998-04-07  
 / PRIOR APPLICATION NUMBER: US 09/223,040  
 / PRIOR FILING DATE: 1998-12-30  
 / NUMBER OF SEQ ID NOS: 46  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO: 22  
 / LENGTH: 600  
 / TYPE: PRT  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / BEAUTY:  
 / OTHER INFORMATION: Description of Artificial Sequence:tri-fusion

US-09-287-849-22  
 Query Match 100.0%; Score 1949; DB: 9; Length: 600;  
 Best Local Similarity 100.0%; Pred. No. 2e-145; Mismatches 0; Indels 0; Gaps 0;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLIVAAQMWDSSAIDLPSAASFQSYWGLTVGSMWG 60  
 Db 9 MVDFGALPPEINSARMYAGPGSASLIVAAQMWDSSAIDLPSAASFQSYWGLTVGSMWG 68

Qy 61 SSAGIMVAASPYVAMSVTAGOAEITAQVRVAAAYETAYGLTVPPVIAENRAELMI 120  
 Db 69 SSAGIMVAASPYVAMSVTAGOAEITAQVRVAAAYETAYGLTVPPVIAENRAELMI 128

Qy 121 LIATNLGQNTPAIAVNEAEYGENWAQDAAMFCAAAATATALLPFEAPEMTSAGG 180  
 Db 129 LIATNLGQNTPAIAVNEAEYGENWAQDAAMFCAAAATATALLPFEAPEMTSAGG 188

Qy 181 LIEQQAAYVEEASDTAANQMLNNYQQLQLAQPTQGTPSSKLGMLWKTVPSPHRSPISN 240  
 Db 189 LIEQQAAYVEEASDTAANQMLNNYQQLQLAQPTQGTPSSKLGMLWKTVPSPHRSPISN 248

Qy 241 MVSMMANNHSMNTNSGYSMTNTLSSMLKGFPAPAAAQQAVCTAAQNGVRAMSIGSSLGSSG 300  
 Db 249 MVSMMANNHSMNTNSGYSMTNTLSSMLKGFPAPAAAQQAVCTAAQNGVRAMSIGSSLGSSG 308

Qy 301 LGGGVANILGRAASVSSLVQAWAANQAVTPARALPLTSAAEPRGPQMLGGLPV 360  
 Db 309 LGGGVANILGRAASVSSLVQAWAANQAVTPARALPLTSAAEPRGPQMLGGLPV 368

Qy 361 GONGARAGGGSGLVLRVPPRYPVMPHSPAG 391  
 Db 369 GONGARAGGGSGLVLRVPPRYPVMPHSPAG 399

RESULT 10  
 US-10-359-460-22  
 ; Sequence 22, Application US/10359460  
 , Sequence 22, Application US/10359460

Publication No. US20030147911A1  
 GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Alderson, Mark  
 APPLICANT: Corixa Corporation  
 TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens and Their Uses  
 FILE REFERENCE: 014058-009081US  
 CURRENT APPLICATION NUMBER: US/10/359,460  
 PRIOR APPLICATION NUMBER: US/09/287,849  
 PRIOR FILING DATE: 1999-04-07  
 PRIOR APPLICATION NUMBER: US 08/818,112  
 PRIOR FILING DATE: 1997-03-13  
 PRIOR FILING DATE: 1998-04-07  
 PRIOR FILING DATE: 1998-12-30  
 NUMBER OF SEQ ID NOS: 46  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 22  
 LENGTH: 600  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence-mutated  
 US-10-369-983-2

Query Match 100.0% Score 1949; DB 15; Length 723;  
 Best Local Similarity 100.0%; Pred. No. 2, 6e-145;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MVDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSASAFOQVWGLTVGSWIG 60  
 Db 333 MVDFGALPPEINSARMYAGPGSASLVAQQMWDVSASSDLSASAFOQVWGLTVGSWIG 392

Query Match 100.0% Score 1949; DB 15; Length 723;  
 Best Local Similarity 100.0%; Pred. No. 2, 6e-145;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 61 SSAGIMVAQASPYAMSTTGAELTAQVRTAAAYBTAXGLTVPPVIAENRAELMI 120  
 Db 393 SSAGIMVAQASPYAMSTTGAELTAQVRTAAAYBTAXGLTVPPVIAENRAELMI 452

Query Match 100.0% Score 1949; DB 15; Length 723;  
 Best Local Similarity 100.0%; Pred. No. 2, 6e-145;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 121 LIATNLIGONTPAIVNAEAYGEMWAQDAAAVGYAAATAATATLIFEEAPEMTSGGG 180  
 Db 453 LIATNLIGONTPAIVNAEAYGEMWAQDAAAVGYAAATAATATLIFEEAPEMTSGGG 512

Query Match 100.0% Score 1949; DB 14; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 2e-145;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0;  
 US-10-359-460-22  
 Qy 1 MVDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSASAFOQVWGLTVGSWIG 60  
 Db 9 MVDFGALPPEINSARMYAGPGSASLVAQQMWDVSASDLSASAFOQVWGLTVGSWIG 68  
 Qy 61 SSAGIMVAQASPYAMSTTGAELTAQVRTAAAYBTAXGLTVPPVIAENRAELMI 120  
 Db 69 SSAGIMVAQASPYAMSTTGAELTAQVRTAAAYBTAXGLTVPPVIAENRAELMI 128  
 Qy 121 LIATNLIGONTPAIVNAEAYGEMWAQDAAAVGYAAATAATATLIFEEAPEMTSGGG 180  
 Db 129 LIATNLIGONTPAIVNAEAYGEMWAQDAAAVGYAAATAATATLIFEEAPEMTSGGG 188  
 Qy 181 LLEQAAAVEEASDTRAANQMLNVPAQLOQLAQPTQGTPSKLGGIWKTVSPHRSPISN 240  
 Db 189 LLEQAAAVEEASDTRAANQMLNVPAQLOQLAQPTQGTPSKLGGIWKTVSPHRSPISN 248  
 Qy 241 MVSMANTHMSNTGSMNTLSSMLKGPAAAAOAQVTAQNYVRANSLSLGSSG 300  
 Db 249 MVSMANTHMSNTGSMNTLSSMLKGPAAAAOAQVTAQNYVRANSLSLGSSG 308  
 Qy 301 LGGGVAAANLGRASVGSGLSPVQPAWAAAANQVTPAARALPLTSLSAERGPQMLGGLPV 360  
 Db 633 LGGGVAAANLGRASVGSGLSPVQPAWAAAANQVTPAARALPLTSLSAERGPQMLGGLPV 692  
 Qy 361 GONGARAGGGLSGLGVRLVPFRPYMPHSPPAG 391  
 Db 693 GONGARAGGGLSGLGVRLVPFRPYMPHSPPAG 723

RESULT 12  
 US-09-886-349A-18  
 Sequence 18, Application US/09886349A  
 Publication No. US2004086523A1  
 GENERAL INFORMATION:  
 APPLICANT: Skeiky, Yasir  
 APPLICANT: Reed, Steven  
 APPLICANT: Alderson, Mark  
 APPLICANT: Corixa Corporation  
 CURRENT FILING DATE: 2001-06-20  
 FILE REFERENCE: 014058-009070US  
 CURRENT APPLICATION NUMBER: US/09/886,349A  
 PRIOR APPLICATION NUMBER: US 09/597,796  
 PRIOR FILING DATE: 2000-06-20  
 PRIOR APPLICATION NUMBER: US 60/265,737  
 PRIOR FILING DATE: 2001-02-01  
 NUMBER OF SEQ ID NOS: 50  
 SOFTWARE: PatentIn Ver. 2.1

RESULT 11  
 US-10-369-983-2  
 Sequence 2, Application US/10369983  
 Publication No. US2003235593A1  
 GENERAL INFORMATION:

SEQ ID NO 18  
 LENGTH: 729  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE: Description of Artificial Sequence:MTB72FMutSA  
 OTHER INFORMATION: (Rai12-TBHP-Ra35MutSA)  
 US-09-886-249A-18

Query Match: 100.0%; Score 1949; DB 12; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 2. 6e-145; Indels 0; Gaps 0;  
 Matches 391; Conservative 0; Mismatches 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLVLAAQMDVSADLFSASAFAQSVMWGLTVGSWIG 60  
 Db 142 MVDFGALPPEINSARMYAGPGSASLVLAAQMDVSADLFSASAFAQSVMWGLTVGSWIG 201

Qy 61 SSAGIMVAASAPYAMSYTAGELTAQVRAAAVETAYGLTPPPVIAENRAELMI 120  
 Db 202 SSAGIMVAASAPYAMSYTAGELTAQVRAAAVETAYGLTPPPVIAENRAELMI 261

Qy 121 LIATNLQONTPAIAYNEAEYGMWAQDAAMFGYAATAATATATLPLFEAPEMTSAGG 180  
 Db 262 LIATNLQONTPAIAYNEAEYGMWAQDAAMFGYAATAATATLPLFEAPEMTSAGG 321

Qy 161 LIEQAAAVEASDTAAANOLMNNPQLQLAQOPTCGTTPSSKLGGLWKTVPSPRSPISN 240  
 Db 322 LIEQAAAVEASDTAAANOLMNNPQLQLAQOPTCGTTPSSKLGGLWKTVPSPRSPISN 381

Qy 241 MVSMANNHNSMTNSGSMTNTLSMLKGAPAAAQVQTAQNGVRAMSILOSSIGSSG 300  
 Db 382 MVSMANNHNSMTNSGSMTNTLSMLKGAPAAAQVQTAQNGVRAMSILOSSIGSSG 441

Qy 301 LGGGVAANIIGRAASVGSLSYPOQAANAOAVTPAARALPLTSLSAERGPQMLGGLPV 360  
 Db 442 LGGGVAANIIGRAASVGSLSYPOQAANAOAVTPAARALPLTSLSAERGPQMLGGLPV 501

Qy 361 GQMGRAGGGISGVLRVPRPPYMPHSPAG 391  
 Db 502 GQMGRAGGGISGVLRVPRPPYMPHSPAG 532

RESULT 14  
 US-10-369-983-21  
 ; Sequence 21, Application US/10369983  
 ; Publication No. US20000235593A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; INVENTION: Guderian, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; CURRENT APPLICATION NUMBER: US10/369,983  
 ; FILE REFERENCE: 014058-003081US  
 ; PRIORITY DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 21  
 ; LENGTH: 729  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:MTB72F  
 US-10-369-983-21

Query Match: 100.0%; Score 1949; DB 15; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-145; Indels 0; Gaps 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLVLAAQMDVSADLFSASAFAQSVMWGLTVGSWIG 60  
 Db 142 MVDFGALPPEINSARMYAGPGSASLVLAAQMDVSADLFSASAFAQSVMWGLTVGSWIG 201

Qy 61 SSAGIMVAASAPYAMSYTAGELTAQVRAAAVETAYGLTPPPVIAENRAELMI 120  
 Db 202 SSAGIMVAASAPYAMSYTAGELTAQVRAAAVETAYGLTPPPVIAENRAELMI 261

Qy 121 LIATNLQONTPAIAYNEAEYGMWAQDAAMFGYAATAATATLPLFEAPEMTSAGG 180  
 Db 262 LIATNLQONTPAIAYNEAEYGMWAQDAAMFGYAATAATATLPLFEAPEMTSAGG 321

Qy 181 LIEQAAAVEASDTAAANOLMNNPQLQLAQOPTCGTTPSSKLGGLWKTVPSPRSPISN 240  
 Db 322 LIEQAAAVEASDTAAANOLMNNPQLQLAQOPTCGTTPSSKLGGLWKTVPSPRSPISN 381

RESULT 15  
 US-10-098-732A-18  
 ; Sequence 18, Application US/10098732A  
 ; GENERAL INFORMATION: No. US20030175294A1  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Brannon, Mark  
 ; APPLICANT: Guderian, Jeffrey  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
 ; FILE REFERENCE: 014058-012201019  
 ; CURRENT APPLICATION NUMBER: US10/098,732A  
 ; PRIOR APPLICATION NUMBER: US 60/275,837  
 ; PRIORITY FILING DATE: 2001-03-13  
 ; NUMBER OF SEQ ID NOS: 80  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 18  
 ; LENGTH: 729  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:MTB72FMutSA  
 ; OTHER INFORMATION: (Rai12-TBHP-Ra35MutSA)  
 US-10-098-732A-18

Query Match: 100.0%; Score 1949; DB 14; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-145; Indels 0; Gaps 0;

Search Completed: August 25, 2004, 03:28:09  
 Job Time : 131 secs

---

RESULT 15  
 US-10-369-983-22  
 / Sequence 22, Application US/10369983  
 ; Publication No. US20030235593A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiry, Yasir  
 ; REED, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 01408-009081US  
 CURRENT APPLICATION NUMBER: US/10/369,983  
 CURRENT FILING DATE: 2003-02-18  
 PRIOR APPLICATION NUMBER: US 60/357,351  
 PRIOR FILING DATE: 2002-02-15  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 22  
 LENGTH: 729  
 / TYPE: PRT  
 / ORGANISM: Artificial Sequence  
 FEATURE:  
 / OTHER INFORMATION: Description of Artificial Sequence: mutated  
 / OTHER INFORMATION: MTb72Fmutsa (Mtbo72F-mutsa)  
 US-10-369-983-22

	Query Match	Score	Length	Best Local Similarity	Pred. No.	Mismatches	Indels	Gaps	O:
Qy	1 MVDGFALPPPEINSARMYTAGPSASLVAAGQWDSVADLFSASAFOSVVGLTVGSWIG	194	145	100.0%	2	6e-145	0	0	
Db	142 MVDGFALPPPEINSARMYTAGPSASLVAAGQWDSVADLFSASAFOSVVGLTVGSWIG	60	201						
Qy	61 SSAGLMVIAASAPVYAMSVTVAQGAELTAQVRAAAAYETAYGLTVPPPVIAENRAELMI	120	120						
Db	202 SSAGLMVIAASAPVAMSVTVAQGAELTAQVRAAAAYETAYGLTVPPPVIAENRAELMI	261	261						
Qy	121 LIATNLGONTPAJAVNEAEYGEMWAQDAAAAMEGYAAATAATATATLPFEAPEMTSAGG	180	180						
Db	262 LIATNLGONTPAJAVNEAEYGEMWAQDAAAAMEGYAAATAATATLPFEAPEMTSAGG	321	321						
Qy	181 LLIQQAEEVEASDTAAANQLMNVPQLQQLAQPTQTTPSSRLGGIWKITVSPHSPISPIN	240	240						
Db	322 LLBQAAVEEASDTAAANQLMNVPQLQQLAQPTQTTPSSKLGGIWKITVSPHSPISPIN	381	381						
Qy	241 MWSMANNHMSMTNIGVSNTNTLSSMLKGFAAAAAQVOTAQNGYRAMSSLGSSLGSSG	300	300						
Db	382 MWSMANNHMSMTNIGVSNTNTLSSMLKGFAAAAAQVQTAQNGYRAMSSLGSSLGSSG	441	441						
Qy	301 LGGGVAANLGRASVGSLSVPOWAANAOVTPAARALPLTSLSAERGPGQMLGGIPV	360	360						
Db	442 LGGGVAANLGRASVGSLSVPOWAANAOVTPAARALPLTSLSAERGPGQMLGGIPV	501	501						
Qy	361 GONGARAGGGSGVLRVPPRPTVMPISPAAG	391	391						
Db	502 GONGARAGGGSGVLRVPPRPTVMPHS PAAG	532	532						

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 25, 2004, 02:57:18 ; Search time 128 Seconds

(without alignments)

863,094 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: MVDFGALPDEINSARMYAGP.....SGVLRVPPRPTMYPHSPAAG 391.

Scoring table: BIOSUM62

Gapext 0.5

Searched: 1586107 seqs., 282547505 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04;\*

1: geneseqP19B03;\*

2: geneseqP19908;\*

3: geneseqP20005;\*

4: geneseqP20015;\*

5: geneseqP20028;\*

6: geneseqP2003as;\*

7: geneseqP2003hs;\*

8: geneseqP2004s;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1949	100.0	391	2	AAW32381		Raw32381 Mycobacte
2	1949	100.0	391	2	AAW32449		Raw32449 Mycobacte
3	1949	100.0	391	2	AAW63335		Raw63335 Mycobacte
4	1949	100.0	391	2	AAWB1702		Raw81702 M. tuberc
5	1949	100.0	391	2	AAVO4778		Raw4778 Mycobacte
6	1949	100.0	391	2	AAV8989		Raw8989 M. tuberc
7	1949	100.0	391	2	AAV39132		Raw39132 M. tuberc
8	1949	100.0	391	4	AAU01888		Aau01888 M. tuberc
9	1949	100.0	391	5	AAE29707		Aee29707 Mycobacte
10	1949	100.0	391	5	Aae11571		Aae11571 Mycobacte
11	1949	100.0	596	2	AAV32070		Raw32070 Mycobacte
12	1949	100.0	596	5	AAE29710		Aae29710 Mycobacte
13	1949	100.0	596	5	AAE15754		Aae15754 Antigenic
14	1949	100.0	599	5	AAU74599		Aau74599 Antigenic
15	1949	100.0	600	2	AAV32068		Raw32068 Mycobacte
16	1949	100.0	600	5	AAU74597		Aau74597 Antigenic
17	1949	100.0	723	7	ADA26354		Ada26354 Mycobacte
18	1949	100.0	729	4	AAO22142		Aao22142 Ra12-H9-3
19	1949	100.0	729	5	AAE29709		Aae29709 Mycobacte
20	1949	100.0	729	5	AAE17513		Aae17513 Mycobacte
21	1949	100.0	729	7	ADA26374		Ada26374 Mycobacte
22	1949	100.0	744	4	AAU01903		Aau01903 M. tuberc
23	1949	100.0	813	7	ADA26367		Ada26367 Mycobacte
24	1949	100.0	815	4	AAU01904		Aau01904 M. tuberc
25	1949	100.0	815	7	ADA26366		Ada26366 Mycobacte

## ALIGNMENTS

RESULT 1  
AAW32381  
ID AAW32381 standard; protein; 391 AA.

XX

AC

XX

DT

13-JAN-1998

(first entry)

XX

DE

Mycobacterium tuberculosis antigen TbH-9FL.

XX

KW

XX

KW

XX

OS

XX

Mycobacterium tuberculosis.

XX

PN

W09709429-A2.

XX

PD

13-MAR-1997.

XX

PF

30-AUG-1996;

96WO-US014675.

XX

PR

01-SEP-1995;

95US-00523435.

PR

22-SEP-1995;

95US-0053136.

PR

22-MAR-1996;

96US-0062280.

PR

05-JUN-1996;

96US-0065800.

PR

12-JUL-1996;

96US-00680573.

XX

(CORIXA CORP.

PA

XX

PR

01-SEP-1995;

95US-0053136.

PR

22-SEP-1995;

95US-0062280.

PR

05-JUN-1996;

96US-0065800.

PR

12-JUL-1996;

96US-00680573.

XX

CORIXA CORP.

PA

XX

Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;

PI

Vedvick TH, Twardzik DR,

PI

XX

DR WPI; 1997-192904/17.

DR N-PSDB; AAT91455.

XX

New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -

PT

useful for diagnosis of M. tuberculosis infection.

XX

Example 3; Page 150-152; 190PP; English.

PS

XX

A new immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its

CC

variant differing only in conservative substitutions and/or

CC

modifications). The present sequence represents a M. tuberculosis antigen,

CC

TbH-9FL. The immunogenic polypeptide can be used to diagnose

CC

M. tuberculosis infection by forming complexes with specific antibodies in

CC

the sample. Fragments of DNA encoding the immunogenic polypeptide can be

CC

used as diagnostic primers or probes and agents that bind to the antigen,

CC

used as diagnostic primers or probes and agents that bind to the antigen,

especially monoclonal antibodies or equivalent polyclonal antibodies, are also used for diagnosis

XX Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1\_2e-142; Indels 0; Gaps 0;  
Matches 391; Conservative 0; Mismatches 0; PT New immunogenic polypeptide(s) from *Mycobacterium tuberculosis* - are useful in vaccines for prevention or treatment of tuberculosis, also for diagnosis.

XX PS Example 3; Page 138-139; 168pp; English.

CC A new immunogenic polypeptide has been developed comprising an immunogenic part of a soluble *Mycobacterium tuberculosis* antigen (or its variant differing only in conservative substitutions and/or modifications). The present sequence represents a *M. tuberculosis* antigen, TbH-9FL. The immunogenic protein, and fusion proteins containing one or more of the proteins plus ESAT-6, are useful in vaccines, preferably when formulated with a non-specific adjuvant, to induce an immune response against *M. tuberculosis* (for treatment or prevention)

SQ XX sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.2e-142; Indels 0; Gaps 0;  
Matches 391; Conservative 0; Mismatches 0; PT New immunogenic polypeptide(s) from *Mycobacterium tuberculosis* - are useful in vaccines for prevention or treatment of tuberculosis, also for diagnosis.

XX PS Example 3; Page 138-139; 168pp; English.

CC A new immunogenic polypeptide has been developed comprising an immunogenic part of a soluble *Mycobacterium tuberculosis* antigen (or its variant differing only in conservative substitutions and/or modifications). The present sequence represents a *M. tuberculosis* antigen, TbH-9FL. The immunogenic protein, and fusion proteins containing one or more of the proteins plus ESAT-6, are useful in vaccines, preferably when formulated with a non-specific adjuvant, to induce an immune response against *M. tuberculosis* (for treatment or prevention)

SQ XX sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.2e-142; Indels 0; Gaps 0;

CC 1 MVDFGALPPEINSARMYAGGSASLVAACQMWDVASDLFSAASAFOSVVGLTVGSMIG 60  
CC 1 MVDFGALPPEINSARMYAGGSASLVAACQMWDVASDLFSAASAFOSVVGLTVGSMIG 60  
CC 61 SSAGIMVAASPSYAMSTAGQELTAQVRVAAAAYETAYGLTVPPVILENRAELMI 120  
CC 61 SSAGIMVAASPSYAMSTAGQELTAQVRVAAAAYETAYGLTVPPVILENRAELMI 120  
CC 121 LIATNLIGQNTPAAVNEAEYGENWAQDAAMTGYAAATAATATLPLFEERPEMTSAGG 180  
Db 121 LIATNLIGQNTPAAVNEAEYGENWAQDAAMTGYAAATAATATLPLFEERPEMTSAGG 180  
Qy 181 LLEQAAVEEASDIAAANQMLNNVPQALQLAQPTQGTTPSSKGIGIWRKTVSPHRSPISN 240  
Db 181 LLEQAAVEEASDIAAANQMLNNVPQALQLAQPTQGTTPSSKGIGIWRKTVSPHRSPISN 240  
Qy 241 MVSNANNINSMTNCSMTNSVMTNLSMKGFAAAAAAQVTAONGVRAMSSLGSSLGSSG 300  
Db 241 MVSNANNINSMTNCSMTNSVMTNLSMKGFAAAAAAQVTAONGVRAMSSLGSSLGSSG 300  
Qy 301 LGGGVAAANTGRAASVGSLSPVQOAWAAANCAVTPAARALPLTSAAERGPQMLGGLPV 360  
Db 301 LGGGVAAANTGRAASVGSLSPVQOAWAAANCAVTPAARALPLTSAAERGPQMLGGLPV 360  
Qy 361 GOMGARAGGGSLGVLRVPPRPPYMPHSAG 391  
Db 361 GOMGARAGGGSLGVLRVPPRPPYMPHSAG 391

RESULT 2  
ID AAN32449 standard; protein; 391 AA.  
XX AC AAN32449;  
XX DT 09-JAN-1998 (first entry)  
XX DE Mycobacterium tuberculosis antigen TbH-9FL.  
XX KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
KW skin testing; M. tuberculosis.  
XX Mycobacterium tuberculosis.  
CS WO9709428-A2.  
XX PD 13-MAR-1997.  
XX PF 30-AUG-1996; 96WO-US014674.  
XX PR 01-SEP-1995; 95US-00523436.  
PR 22-SEP-1995; 95US-00533634.  
PR 22-MAR-1996; 96US-00620874.  
PR 05-JUN-1996; 96US-00659633.  
PR 12-JUL-1996; 96US-00680574.  
XX (CORTI-) CORIXA CORP.  
PA Reed SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;  
PI vedwick TH, Twardzik DR;  
XX WPI: 1997-192903/17.  
DR N-PSDB; AAT191521.  
XX

XX PT 17-OCT-2003 (revised)

XX DT 09-NOV-1998 (first entry)

XX DE Mycobacterium tuberculosis antigen TbH-9FL.

XX KW Tuberculosis; infection; diagnosis; antigen; TbH-9FL.

XX OS Mycobacterium tuberculosis; strain H37Rv.

XX PN WO9816645-A2.

XX PD 23-APR-1998.

XX PP 07-OCT-1997; 97WO-US018214.

XX PR 11-OCT-1996; 96US-00729622.

PR	13-MAR-1997;	97US-00818111.	DE	M. tuberculosis immunogenic polypeptide TbH-9FL.
XX	(CORTI-)	CORIXA CORP.	XX	Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
PA			KW	vaccine; pharmaceutical; infection; diagnosis.
XX	Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;		KW	
PI	PI		XX	
PI	Vedwick TS, Twardzik DR, Lodes MJ;		OS	Mycobacterium tuberculosis.
XX			XX	
DR	WPI: 1998-251292/22.		XX	WO9816646-A2.
DR	N-PSDB;	AAV4395.	PN	
XX			PD	23 - ABR-1998.
PT	New isolated Mycobacterium tuberculosis polypeptides and DNA - used to develop products for the detection of M. tuberculosis infection and diagnosis of tuberculosis.		XX	97WO-US018293.
PT			PF	07-OCT-1997;
PT			XX	96US-00730510.
XX			PR	11-OCT-1996;
PS	PS	Example 3; Page 133-135; 250pp; English.	PR	13-MAR-1997;
XX			PR	97US-00816112.
XX			XX	XX (CORTI-) CORIXA CORP.
CC	This polypeptide comprises Mycobacterium tuberculosis antigen TbH-9FL. It is encoded by genomic DNA (see AAV4395) isolated from a M. tuberculosis strain F37Rv genomic library using a probe from clone TbH-9 (see AAV4371). The invention relates to compositions and methods for diagnosing tuberculosis. It provides a polypeptide of a soluble M. tuberculosis antigen, comprising an antigenic portion of an M. tuberculosis antigen, as well as DNA sequences encoding such polypeptides, recombinant expression vectors and transformed or transfected host cells. Also claimed are methods and diagnostic kits for detecting M. tuberculosis infection in a patient using these polypeptides, antibodies or oligonucleotide probes and primers for the diagnosis of tuberculosis. (Updated on 17-Oct-2003 to standardise OS field)		XX	XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
CC			PI	PI Vedwick TS, Twardzik DR, Lodes MJ;
CC			XX	XX WPI: 1998-261042/23.
CC			DR	DR N-PSDB; AAV64503.
CC			XX	XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to develop products for the detection of M. tuberculosis infection and for diagnosis, treatment and prevention of tuberculosis.
CC			XX	XX PS Example 3B; Page 128-129; 230pp; English.
CC			XX	XX This sequence represents an immunogenic portion of a soluble Mycobacterium tuberculosis (Mt) antigen which can be used in a method for inducing protective immunity against tuberculosis (TB). This sequence can be formulated into vaccines and/or pharmaceutical compositions for immunising against M. tuberculosis infection or may be used for the diagnosis of tuberculosis.
CC			XX	XX Sequence 391 AA;
CC			XX	XX Query Match Score 1949; DB 2; Length 391; Best Local Similarity 100.0%; Prod. No. 1.2e-142; Mismatches 0; Indels 0; Gaps 0; Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
SQ	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60	Qy	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60
Db	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60	Db	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60
Qy	61 SSAGIMVAASPYAAMSTAGQELTAQVRAAAAYETAYGLITVPPVIAENRAELM 120	61 SSAGIMVAASPYAAMSTAGQELTAQVRAAAAYETAYGLITVPPVIAENRAELM 120	Qy	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60
Db	61 SSAGIMVAASPYAAMSTAGQELTAQVRAAAAYETAYGLITVPPVIAENRAELM 120	61 SSAGIMVAASPYAAMSTAGQELTAQVRAAAAYETAYGLITVPPVIAENRAELM 120	Db	1 MVDFGALPPPEINSARMYAGPGSASLVAAMQMWDVASDIFSAASAQSVMWGLTVGWMIG 60
Qy	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180	Qy	61 SSAGIMVAASPYAAMSTAGQELTAQVRAAAAYETAYGLITVPPVIAENRAELM 120
Db	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180	Db	61 SSAGIMVAASPYAAMSTAGQELTAQVRAAAAYETAYGLITVPPVIAENRAELM 120
Qy	181 LLEQQAAVEEASDITAANQLMNNTGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 240	181 LLEQQAAVEEASDITAANQLMNNTGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 240	Qy	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180
Db	181 LLEQQAAVEEASDITAANQLMNNTGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 240	181 LLEQQAAVEEASDITAANQLMNNTGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 240	Db	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180
Qy	241 MVSMMNHMSMTNSGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 300	241 MVSMMNHMSMTNSGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 300	Qy	181 LLEQQAAVEEASDITAANQLMNNTGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 240
Db	241 MVSMMNHMSMTNSGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 300	241 MVSMMNHMSMTNSGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 300	Db	121 LIATNLIGQNTPAIAVNEAEYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180
Qy	301 LGGGVAAANLGRASAASVGSLSYQPMQAANQAVTPAIALPITSLTSAEAGPGONLGGLPV 360	301 LGGGVAAANLGRASAASVGSLSYQPMQAANQAVTPAIALPITSLTSAEAGPGONLGGLPV 360	Qy	241 MVSMMNHMSMTNSGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 300
Db	301 LGGGVAAANLGRASAASVGSLSYQPMQAANQAVTPAIALPITSLTSAEAGPGONLGGLPV 360	301 LGGGVAAANLGRASAASVGSLSYQPMQAANQAVTPAIALPITSLTSAEAGPGONLGGLPV 360	Db	241 MVSMMNHMSMTNSGSMINTLSSMLKGFPAAQVTAQNGVRANSLSIGSG 300
Qy	361 GOMGARAGGGLGSGVLRVPREPYMPHS PAAG 391	361 GOMGARAGGGLGSGVLRVPREPYMPHS PAAG 391	Qy	361 GOMGARAGGGLGSGVLRVPREPYMPHS PAAG 391
Db	361 GOMGARAGGGLGSGVLRVPREPYMPHS PAAG 391	361 GOMGARAGGGLGSGVLRVPREPYMPHS PAAG 391	Db	361 GOMGARAGGGLGSGVLRVPREPYMPHS PAAG 391

RESULT 4

AAW81702

ID AAW81702 standard; protein; 391 AA.

XX

AC AAW81702;

DT 27-JAN-1999 (first entry)

RESULT 5

QY	361	GQMGRAGGGLSCGVLRVPPRPPYMPHSPAG	391			
ID	361	GQMGRAGGGLSCGVLRVPPRPPYMPHSPAG	391			
XX						
AC						
XX						
DT	06-JUL-1999	(first entry)				
XX						
XX						
DE						
XX						
KW						
KW						
KW						
OS						
XX						
Mycobacterium sp.						
PN	W0909186-A2.					
XX						
PR	14-AUG-1998;	98WO-FR001813.				
XX						
PR	14-AUG-1997;	97FR-00010404.				
XX						
PR	11-SEP-1997;	97FR-00011325.				
XX						
PA	(INSP ) INST PASTEUR.					
XX						
PI	Gicquel B,	Portnoi D,	Lim E,	Pelicic V,	Guigueno A;	
PI	Goguet De La Salmoniere Y;					
XX						
DR	1999-181045/15.					
DR	N-PSDB; AX34030.					
XX						
PT	Mycobacterial DNA vectors containing reporter constructs - for identifying coding or promoter sequences involved in infection-associated protein expression.					
PT	Claim 32; Fig 5R; 309pp; French.					
XX						
CC	Sequences AAY04742-Y05000 and AAY07201-Y07204 represent secreted proteins from various Mycobacterium species microorganisms. The encoding nucleotide sequences can be used as primers and probes for methods for detecting and identifying mycobacteria, especially belonging to the M. tuberculosis complex. The encoded proteins can be used in vaccines for immunisation against a bacterial or viral infection					
CC	Sequence 391 AA;					
CC	Query Match 100.0%; Score 1949; DB 2;	Length 391;				
CC	Best Local Similarity 100.0%; Pred. No. 1.2e-142;					
CC	Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
XX						
QY	1	MDFGALPPPEINSARMYAGPGSASLYAAQWMDSVASDLFSAASAFQVVAAYETAYGLTVGSGW	60			
Db	1	MDFGALPPPEINSARMYAGPGSASLYAAQWMDSVASDLFSAASAFQVVAAYETAYGLTVGSGW	60			
QY	61	SSAGLVAAYAASPVAVMSVTAGQAEIYTAQYRVAAAYETAYGLTVGSGW	60			
Db	61	SSAGLVAAYAASPVAVMSVTAGQAEIYTAQYRVAAAYETAYGLTVGSGW	60			
QY	121	LATNLIGONTPAIAVNEAEYGEMMAODAAAMEGYAAATAATATLPPBAPMTSAGG	180			
Db	121	LATNLIGONTPAIAVNEAEYGEMMAODAAAMEGYAAATAATATLPPBAPMTSAGG	180			
QY	121	LATNLIGONTPAIAVNEAEYGEMMAODAAAMEFYAAATAATATLPPBAPMTSAGG	180			
Db	121	LATNLIGONTPAIAVNEAEYGEMMAODAAAMEFYAAATAATATLPPBAPMTSAGG	180			
QY	181	ILEQAAVEEASDTAANQLNNVPOLQQAQOPTGTTPSKLGILWKTVSPHSPISN	240			
Db	181	ILEQAAVEEASDTAANQLNNVPOLQQAQOPTGTTPSKLGILWKTVSPHSPISN	240			
QY	241	MVSMAHHMSTMNTSGVSMNTTSSMLKGFAAAAQAVQTAQNGYRAMSSLGSSG	300			
Db	241	MVSMAHHMSTMNTSGVSMNTTSSMLKGFAAAAQAVQTAQNGYRAMSSLGSSG	300			
QY	301	IGGGTAANLGRAASVGSLSVCAWAAANOATTPAARALPLTSLSAERGGCQMLGLPV	360			
Db	301	IGGGTAANLGRAASVGSLSVCAWAAANOATTPAARALPLTSLSAERGGCQMLGLPV	360			
QY	301	IGGGTAANLGRAASVGSLSVCAWAAANOATTPAARALPLTSLSAERGGCQMLGLPV	360			
Db	301	IGGGTAANLGRAASVGSLSVCAWAAANOATTPAARALPLTSLSAERGGCQMLGLPV	360			
QY	241	MVSMAHHMSTMNTSGVSMNTTSSMLKGFAAAAQAVQTAQNGYRAMSSLGSSG	300			
Db	241	MVSMAHHMSTMNTSGVSMNTTSSMLKGFAAAAQAVQTAQNGYRAMSSLGSSG	300			
QY	181	LLEQAAVEEASDTAANQLNNVPOLQQAQOPTGTTPSKLGILWKTVSPHSPISN	240			
Db	181	LLEQAAVEEASDTAANQLNNVPOLQQAQOPTGTTPSKLGILWKTVSPHSPISN	240			
QY	241	MVSMAHHMSTMNTSGVSMNTTSSMLKGFAAAAQAVQTAQNGYRAMSSLGSSG	300			
Db	241	MVSMAHHMSTMNTSGVSMNTTSSMLKGFAAAAQAVQTAQNGYRAMSSLGSSG	300			

Db 241 NYSMANNHMSMTNSGTMNTLSSMLKGFAAPAAQAVQTAQNGTRAMSGLGSSLGSSG 300 Db 61 SSAGIIVAAASPYAMSVTAGAELTAQVRVAAAAYETATGTLTVPPVIAENRAELMI 120 Qy 301 LGGVAANLGRAASVGSLSPDAWALANQAVTPAARALPLTSITSAERSPGOMLGGLPV 360 Qy 301 LGGVAANLGRAASVGSLSPDAWALANQAVTPAARALPLTSITSAERSPGOMLGGLPV 360 Db 121 LIATNLIGONTPALAVNEAEYGENMWADAAANGYAAATAATATLPLFEAPEMTSAGG 180 Db 121 LIATNLIGONTPALAVNEAEYGENMWADAAANGYAAATAATATLPLFEAPEMTSAGG 180 Qy 181 LLQQAAVVEASDAAANOLMNVPOQLAQOPTGOTTPSSKLGGLMKTVSPHRSPDSN 240 Qy 181 LLQQAAVVEASDAAANOLMNVPOQLAQOPTGOTTPSSKLGGLMKTVSPHRSPDSN 240 Db 241 MVSMANNHMSMTNSGVMNTLSSMLKGFAAPAAQAVQTAQNGTRAMSGLGSSG 300 Db 241 MVSMANNHMSMTNSGVMNTLSSMLKGFAAPAAQAVQTAQNGTRAMSGLGSSG 300 Qy 241 MVSMANNHMSMTNSGVMNTLSSMLKGFAAPAAQAVQTAQNGTRAMSGLGSSG 300 Qy 181 LLQQAAVVEASDAAANOLMNVPOQLAQOPTGOTTPSSKLGGLMKTVSPHRSPDSN 240 Qy 241 MVSMANNHMSMTNSGVMNTLSSMLKGFAAPAAQAVQTAQNGTRAMSGLGSSG 300 Qy 301 LGGGVAAANLGRAASVGSLSYQAMAAQAVTPAARALPLTSITSAERGPQMLGGCPV 360 Qy 301 LGGGVAAANLGRAASVGSLSYQAMAAQAVTPAARALPLTSITSAERGPQMLGGCPV 360 Db 361 GOMGARGGGLSCGYLRVPPRPMHPSPAG 391 Db 361 GOMGARGGGLSCGYLRVPPRPMHPSPAG 391 Db 361 GOMGARGGGLSCGYLRVPPRPMHPSPAG 391

RESULT 7  
AAV39112 ID AAY39132 standard; protein; 391 AA.  
XX AC AAY39132;  
XX DT 05-NOV-1999 (first entry)  
XX DE M. tuberculosis antigen TbH-9FL amino acid sequence.  
XX DE Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;  
KW immune response; diagnosis; immunisation; vaccine; infection;  
KW Mycobacterium tuberculosis; Mycobacterium; M. tuberculosis; antigen; immunogen;  
KW immune response; skin test.  
XX OS WO9942076-A2.  
XX PN AAU01888  
XX ID AAU01888 standard; protein; 391 AA.  
XX AC AAU01888;  
XX DT 29-AUG-2001 (first entry)  
XX DE M. tuberculosis antigen TbH9 (MtB39A).  
XX DE TbH9; MtB39A; antigen; vaccine; tuberculosis; AIDS;  
KW acquired immunodeficiency disease.  
XX OS Mycobacterium tuberculosis.  
XX PN WO200124820-A1.  
XX PD 12-APR-2001.  
XX PR 10-OCT-2000; 2000WO-US028095.  
XX PR 07-OCT-1999; 99US-0158338P.  
XX PR 07-OCT-1999; 99US-0158425P.  
XX PA (CORIXA CORP.)  
XX PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
PI Vediwick TS, Twardzik DR, Lodes MJ, Hendrickson RC;  
XX DR 1999-527409/44.  
DR N-PSDB; AAZ19105.  
XX PT New antigens from Mycobacterium tuberculosis useful in diagnostic skin  
tests and protective or therapeutic vaccines or compositions.  
XX PS Example 3: Page 123-124; 29pp; English.  
XX CC The present invention describes polypeptides comprising an immunogenic  
part of a Mycobacterium tuberculosis antigen (Ag). Also described are  
vaccines and fusion protein containing M. tuberculosis Ag's. M.  
tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other  
polypeptides fragments can be used in pharmaceutical compositions or  
vaccines to generate a protective or therapeutic immune response to M.  
tuberculosis or as reagents in skin tests for diagnosis of tuberculosis.  
Ag can induce proliferation of, or cytokine secretion by, T, B or natural  
killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249  
to AAZ19450 and AAY39252 to AAY39253 are used in the exemplification of  
the present invention.  
XX SQ Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.2e-142; Mismatches 0; Inels 0; Gaps 0;  
Matches 391; Conservative 0; CC

Qy 1 MDVFGALPPEINSARMYAGPCGSASLYAAQMWDVSASDLSAASFQSTVWGLTVGSNG 60  
Db 1 MDVFGALPPEINSARMYAGPCGSASLYAAQMWDVSASDLSAASFQSTVWGLTVGSNG 60  
Qy 61 SSAGIIVAAASPYAMSVTAGAELTAQVRVAAAAYETATGTLTVPPVIAENRAELMI 120  
Db 61 SSAGIIVAAASPYAMSVTAGAELTAQVRVAAAAYETATGTLTVPPVIAENRAELMI 120

Query Match 100.0%; Score 1949; DB 4; Length 391;

SQ Sequence 391 AA;

Best Local Similarity	100.0%	Pred. No.	1.2e-142;
Matches	391	Conservative	0;
		Mismatches	0;
		Indels	0;
		Gaps	0;
Dy	1	MVDFGALPPEINSARMYAGCOSASLYAAQNMDSVASYASDLSAASAQFQSYVWGLITVGSMIG	60
Dy	1	MVDFGALPPEINSARMYAGCOSASLYAAQNMDSVASYASDLSAASAQFQSYVWGLITVGSMIG	60
Dy	61	SSAGLMVAAAASPYVAMSVTAGQALTAQVRAAAAYETAYGLTPPEVIAENRAELMI	120
Dy	61	SSAGLMVAAAASPYVAMSVTAGQALTAQVRAAAAYETAYGLTPPEVIAENRAELMI	120
Dy	121	LJATNLGQNTPAIAVNNEAEYGMAQDAAAAMEGYAAATAATATLIPPEEEAPEMTSAGG	180
Dy	121	LJATNLGQNTPAIAVNNEAEYGMAQDAAAAMEGYAAATAATLIPPEEEAPEMTSAGG	180
b	181	LLEQAAVEEASDTAAANQMLNNTYEAQQLQLAQPTQGTTPSKSLKGMLKTVSPHRSPISN	240
b	181	LLEQAAVEEASDTAAANQMLNNTYEAQQLQLAQPTQGTTPSKSLKGMLKTVSPHRSPISN	240
b	241	MVSMANNHMSMTNSGSVMNTLSSMLKGFPAAAAAQVOTAANGVRAMSSLGSSLGSSG	300
b	241	MVSMANNHMSMTNSGSVMNTLSSMLKGFPAAAAAQVOTAANGVRAMSSLGSSLGSSG	300
b	301	LGGGAANLGRAASVGSSLVPOAAANAOAVTPAARALPLTSLSAAERGPQOMTGGLPV	360
b	301	LGGGAANLGRAASVGSSLVPOAAANAOAVTPAARALPLTSLSAAERGPQOMTGGLPV	360
y	361	GQMGRAGGGLSGVLRPVRPYNMPHSPAG	391
b	361	GQMGRAGGGLSGVLRPVRPYNMPHSPAG	391

**RESULT 9**  
 AE22707  
 D AAE29707 standard; protein: 391 AA.  
 X  
 C AAE29707;  
 X  
 X 27-JAN-2003 (first entry)  
 T Mycobacterium sp. TbH9FL antigenic protein.  
 X  
 X Vaccine: immunity; diagnostic agent; gene therapy; TbH9FL antigen.  
 W  
 W

X WO0027792-A2.  
X 19-SEP-2002.  
X 13-MAR-2002; 2002WO-US008223.  
X 13-MAR-2001; 2001US-0275837P.  
X (CORI-) CORIXA CORP.  
X Skeiky Y, Brannon M, Guderian J;

New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF, M15 or 6H Polynucleotide, useful as vaccine to elicit protective immunity against pathogenic microorganisms e.g. Leishmania and Mycobacterium tuberculosis.

Disclosure; Page 86-87; 155pp; English.

The invention relates to a recombinant nucleic acid molecule encoding a fusion polypeptide. The recombinant nucleic acid comprises a heterologous polynucleotide sequence encoding an antigen or an antigenic fragment from *Mycobacterium* sp. and a *Leishmania* polynucleotide sequence encoding a polypeptide or its fragment. The *Leishmania* polynucleotide is selected from N-PSDB; AAD47082.

from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention	SQ	Sequence 391 AA;
CC are used in methods for eliciting immune response in mammals. They are	Query Match	100 %; Score 1949; DB 5; Length 391;
CC used as vaccines to elicit protective immunity against pathogenic	Best Local Similarity	100 %; Pred. No. 1.2e-142;
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion	Matches	391; Mismatches 0; Indels 0; Gaps 0;
CC polypeptides are used for enhancing the expression of polynucleotides, as	Qy	1 MYDFGALPPEINSARYTAGPSASLIVAAQNWDSYASDLSAASAFAFQSYYVWGLTVGSWIG 60
CC in vivo diagnostic agents and for raising antibodies in a non-human	Db	1 MYDFGALPPEINSARYTAGPSASLIVAAQNWDSYASDLSAASAFAFQSYYVWGLTVGSWIG 60
CC animal. The invention is used in gene therapy. The present sequence is	Qy	61 SSAGLMVAAASPYVAMSVTAGQAETTAQVRVAAAYETAYGTTPPPVIAENRAEIMI 120
CC Mycobacterium sp. TbH9FL antigenic protein	Db	61 SSAGLMVAAASPYVAMSVTAGQAETTAQVRVAAAYETAYGTTPPPVIAENRAEIMI 120
XX	Qy	121 LIATNLIGQNTPAIAINNEADYCEMHNQDAAMAFGTYAAATAATATLPLPEEAAPMTSAGG 180
	Db	121 LIATNLIGQNTPAIAINNEADYCEMHNQDAAMAFGTYAAATAATATLPLPEEAAPMTSAGG 180
	Qy	181 LIEQAAVVEEASDTAAQNLMNNVYQALQQLAQPQQTGPSSKLUGLWKTVPSPRSPISN 240
	Db	181 LIEQAAVVEEASDTAAQNLMNNVYQALQQLAQPQQTGPSSKLUGLWKTVPSPRSPISN 240
	Qy	241 MVSMAANTHMSMTNSGYSMTNTLSSMKKGAPAAAQAVOTAQNGVRAMSLGSSLGSSG 300
	Db	241 MVSMAANTHMSMTNSGYSMTNTLSSMKKGAPAAAQAVOTAQNGVRAMSLGSSLGSSG 300
	Qy	301 LGGGVANLGRAASIGSLSVSPQAVAAQANDAVTPAARALPLTSLSAERPGOMLGCPV 360
	Db	301 LGGGVANLGRAASIGSLSVSPQAVAAQANDAVTPAARALPLTSLSAERPGOMLGCPV 360
	Qy	361 GOMGARAGGGLSCHLGRVPPRPYMPHSPPAG 391
	Db	361 GOMGARAGGGLSCHLGRVPPRPYMPHSPPAG 391

DR N-PSDB; AAD28341.  
 XX Composition comprising MTB39 antigen and MTB32A antigen from  
 PT Mycobacterium species, useful for eliciting immune response in a subject.  
 XX  
 PS Claim 83; Page 102-103; 136pp; English.

The present invention relates to fusion proteins containing at least two Mycobacterium species antigens, nucleotides encoding them and compositions comprising such fusion proteins. The present invention particularly relates to nucleic acids encoding fusion proteins that include two or more individual M. tuberculosis antigens which increase the serological sensitivity of sera from individuals infected with tuberculosis and methods for their use in diagnosis, prevention and treatment of tuberculosis infection. Sequences of the invention are useful for eliciting an immune response in a mammal, e.g., human, immunised with BCG. They are useful in the diagnosis, treatment and prevention of Mycobacterium infection. The fusion proteins and the polynucleotides are useful as diagnostic tools in patients infected with Mycobacterium, in vitro and in vivo assays for detecting humoral antibodies or cell-mediated immunity against M. tuberculosis, for the diagnosis of an infection or monitoring of disease progression, as immunogens to generate or elicit a protective immune response in a patient and for raising anti-M. tuberculosis antibodies in a non-human animal. Sequences of the invention are also used as vaccines. MTB32A fusion proteins of the invention are useful as in vivo diagnostic agents for intradermal skin test. The present sequence is Mycobacterium species MTB39 (Tbh9) protein.

XX Sequence 391 AA:  
 SQ Query Match 100.0%; Score 1949; DB 5; Length 391;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFGALPBEINSARMYAGPGASLIVNAQMDSVASDPLSASAFAOSVNVGLTVGSWIG 60  
 Db 1 MVDFGALPBEINSARMYAGPGASLIVNAQMDSVASDPLSASAFAOSVNVGLTVGSWIG 60  
 QY 61 SSALMVAAASPYAMSVTTAGQAELTIAQVRVAAAYETAYGLTVPPVTAENRAELMI 120  
 Db 61 SSAGLMVAAASPYAMSVTTAGQAELTIAQVRVAAAYETAYGLTVPPVTAENRAELMI 120  
 QY 121 LIATNLIGQNTPAIAVNEBYGENWAQDAAMGSAANTATAATLPLFEEPEMTSAGG 180  
 Db 121 LIATNLIGQNTPAIAVNELEYGENWAQDAAMGSAANTATAATLPLFEEPEMTSAGG 180  
 QY 181 LLEQAAAVEBEASDTAAANQMNVPQALQOLAQFTOGTTPSSKLGIGLWKTSPHRSPIN 240  
 Db 181 LLEQAAAVEBEASDTAAANQMNVPQALQOLAQFTOGTTPSSKLGIGLWKTSPHRSPIN 240  
 QY 241 MVSMMNNHHSMTNSGVSMNTLSSMLKGFPAPAAAQQAYQTAAGNGVRAMSSLGSSG 300  
 Db 241 MVSMMNNHHSMTNSGVSMNTLSSMLKGFPAPAAAQQAYQTAAGNGVRAMSSLGSSG 300  
 QY 301 LGGGYAANLGRAASYGSISVPOQAWAAANGQAVTAAPARALPLTSLSAARRGPQGMIGGLPV 360  
 Db 301 LGGGYAANLGRAASYGSISVPOQAWAAANGQAVTAAPARALPLTSLSAARRGPQGMIGGLPV 360  
 QY 361 GQMGRAGGGSLGVLRVPPRYPNMPHSPAG 391  
 Db 361 GQMGRAGGGSLGVLRVPPRYPNMPHSPAG 391

RESULT 11  
 AAY3-2070 ID AAY3-2070 standard; protein; 596 AA.  
 XX AC AAY3-2070;  
 XX DT 17-JAN-2000 (first entry)  
 XX DE Mycobacterium tuberculosis antigen fusion protein MtB59f.

XX KW Tuberculosis; antigen; fusion protein; MtB59f; Tbh9; Ra35; diagnosis;  
 XX KW therapy; vaccine; immunogen.  
 OS XX Mycobacterium tuberculosis.  
 XX Location/Qualifiers  
 KEY 1-8  
 Peptide /note= "Met/His tag"  
 FH 9-140  
 FT Protein /note= "Ra12"  
 FT Protein /note= "Tbh9"  
 FT Protein /note= "Ra12"  
 XX PN WO951748-A2.  
 XX PD 14-OCT-1999.  
 XX PF 07-APR-1999; 99WO-US00717.  
 XX PR 07-APR-1998; 98US-00056556.  
 XX PR 30-DEC-1998; 98US-00223040.  
 XX PA (CORYX CORP.  
 PI Skeity YW, Alderson M, Campos-Neto A;  
 XX DR 1999-601610/51.  
 DR N-PSDB; AAZ2005.  
 XX DR 1999-601610/51.  
 DR N-PSDB; AAZ2005.  
 XX PT New fusion proteins useful for diagnosis, prevention and treatment of tuberculosis.  
 XX PT  
 PS Claim 1; Fig 12A-B; 83pp; English.  
 XX  
 This sequence represents a recombinant Mycobacterium tuberculosis bi-antigen fusion protein, termed MtB59f, composed of the antigens Tbh9 and Ra35. The fusion protein is expressed in host cells using a vector carrying a polynucleotide (see AAY3-2070) comprising the coding sequences for the 2 antigens. The invention provides fusion proteins (see AAY3-2070) containing at least 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides encoding them are useful as vaccines for preventing tuberculosis (claimed), for diagnosis (via in vitro assays or intradermal skin tests for detection of anti-M. tuberculosis antibodies), monitoring of disease progression, and treatment of tuberculosis. They are more effective immunogens than mixtures of the individual protein components  
 XX Sequence 596 AA:  
 SQ Query Match 100.0%; Score 1949; DB 2; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVDGFALPBEINSARMYAGPGASLIVNAQMDSVASDPLSASAFAOSVNVGLTVGSWIG 60  
 Db 9 MVDGFALPBEINSARMYAGPGASLIVNAQMDSVASDPLSASAFAOSVNVGLTVGSWIG 60  
 QY 61 SSAGLMVAAASPYAMSVTTAGQAELTIAQVRVAAAYETAYGLTVPPVTAENRAELMI 120  
 Db 69 SSAGLMVAAASPYAMSVTTAGQAELTIAQVRVAAAYETAYGLTVPPVTAENRAELMI 120  
 QY 121 LIATNLIGQNTPAIAVNEBYGENWAQDAAMGSAANTATAATLPLFEEPEMTSAGG 180  
 Db 121 LIATNLIGQNTPAIAVNELEYGENWAQDAAMGSAANTATAATLPLFEEPEMTSAGG 180  
 QY 129 LIATNLIGQNTPAIAVNEAEQMGYAAATATAATLPLFEEPEMTSAGG 188  
 Db 181 LIQEAAAVEBEASDTAAANQMNVPQALQOLAQFTOGTTPSSKLGIGLWKTSPHRSPIN 240  
 Db 189 LIQEAAAVEBEASDTAAANQMNVPQALQOLAQFTOGTTPSSKLGIGLWKTSPHRSPIN 248  
 QY 241 MVSMMNNHHSMTNSGVSMNTLSSMLKGFPAPAAAQQAYQTAAGNGVRAMSSLGSSG 300  
 Db 249 MVSMMNNHHSMTNSGVSMNTLSSMLKGFPAPAAAQQAYQTAAGNGVRAMSSLGSSG 300





Query Match 100.0%; Score 1949; DB 2; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-142;  
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPBEINSARMYAGPGSASLYAAQMDVSADLFFRAASAFOSUMLTIVGSWIG 60  
 Db 9 MVDFGALPBEINSARMYAGPGSASLYAAQMDVSADLFFRAASAFOSUMLTIVGSWIG 68  
 Qy 61 SSGGLMYAASSPVAVMSVTQAELTAQVRVAAYETPGVPPVTAENAEALMI 120  
 Db 69 SSGGLMYAASSPVAVMSVTQAELTAQVRVAAYETPGVPPVTAENAEALMI 128  
 Qy 121 LIATNLIGQNTPAIAVNEAYGEMWAQDAAMFGTAATAATATLPPFEAPENTSAGG 180  
 Db 129 LIATNLIGQNTPAIAVNEAYGEMWAQDAAMFGTAATAATATLPPFEAPENTSAGG 188  
 Qy 181 LLEQQAAVEEASTAAANOLNNVPQALQOQLAQOPTGTTPSKUGLWKTPSPHSRSPIN 240  
 Db 189 LLEQQAAVEEADTAANQLNNVPQALQOQLAQOPTGTTPSKUGLWKTPSPHSRSPIN 248  
 Qy 241 MVSAMNMHMSMTNSGVSMTNTLSMKGFAAAAQAVTAQNGYRAMSSLGSSG 300  
 Db 249 MVSAMNMHMSMTNSGVSMTNTLSMKGFAAAAQAVTAQNGYRAMSSLGSSG 308

Sequence 600 AA;

XX

PT New fusion proteins useful for diagnosis, prevention and treatment of tuberculosis.

XX

PS Claim 1; Fig 10A-B; 83pp; English.

XX

CC This sequence represents a recombinant Mycobacterium tuberculosis tri-  
 CC antigen fusion protein, termed Mt61f, composed of the antigens TbH9, DPV  
 CC and MTI. The fusion protein is expressed in host cells using a vector  
 CC carrying a polynucleotide (see AAZ2003) comprising the coding sequences  
 CC for the 3 antigens. The invention provides fusion proteins (see AAY32059-  
 CC 71) containing at least 2 M. tuberculosis antigens. The new fusion  
 CC proteins and polynucleotides encoding them are useful as vaccines for  
 CC preventing tuberculosis (claimed), for diagnosis (via in vitro assays or  
 CC intradermal skin tests for detection of anti-M. tuberculosis antibodies),  
 CC monitoring of disease progression, and treatment of tuberculosis. They  
 CC are more effective immunogens than mixtures of the individual protein  
 CC components.

XX

PT Tuberculosis.

XX

PS Skeiky YAW, Alderson M, Campos-Neto A;

XX

DR WPI: 1999-601610/51.

XX

DR N-PSDB; AAZ2003.

XX

PT Tuberculosis.

XX

PS Claim 1; Fig 10A-B; 83pp; English.

XX

CC This sequence represents a recombinant Mycobacterium tuberculosis tri-  
 CC antigen fusion protein, termed Mt61f, composed of the antigens TbH9, DPV  
 CC and MTI. The fusion protein is expressed in host cells using a vector  
 CC carrying a polynucleotide (see AAZ2003) comprising the coding sequences  
 CC for the 3 antigens. The invention provides fusion proteins (see AAY32059-  
 CC 71) containing at least 2 M. tuberculosis antigens. The new fusion  
 CC proteins and polynucleotides encoding them are useful as vaccines for  
 CC preventing tuberculosis (claimed), for diagnosis (via in vitro assays or  
 CC intradermal skin tests for detection of anti-M. tuberculosis antibodies),  
 CC monitoring of disease progression, and treatment of tuberculosis. They  
 CC are more effective immunogens than mixtures of the individual protein  
 CC components.

XX

Qy 301 LGGGYAANLGRAASVGSLSYQPQAWAAANQAVTPARALPLTSAAERGPQMLGGIPV 360  
 Db 309 LGGGYAANLGRAASVGSLSYQPQAWAAANQAVTPARALPLTSAAERGPQMLGGIPV 368

Qy 361 GQM GARAGGGLSGVLRVPFPYVMFHSPIAG 391  
 Db 369 GQM GARAGGGLSGVLRVPFPYVMFHSPIAG 399

Search completed: August 25, 2004,  
 Job time : 131 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 25, 2004, 03:08:19 ; Search time 34 Seconds  
(without alignments)

593,699 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDFGALPPEINSARMYGP.....SGVILRVPFRPYWMPHSPAAAG 391

Scoring table: BLUSON62

Gapext 10.0 , Gapext 0.5

Searched: 389414 seqs., 51625971 residues

Total number of hits satisfying chosen parameters:

389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cn2\_6\_ptodata/2/iaa/5A\_COMB.pep:/\*  
 2: /cn2\_6\_ptodata/2/iaa/5B\_COMB.pep:/\*  
 3: /cn2\_6\_ptodata/2/iaa/6A\_COMB.pep:/\*  
 4: /cn2\_6\_ptodata/2/iaa/6B\_COMB.pep:/\*  
 5: /cn2\_6\_ptodata/2/iaa/FACTUS\_COMB.pep:/\*  
 6: /cn2\_6\_ptodata/2/iaa/backfile.pep:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1949	100.0	391	3	US-09-818-112-107- <del>20</del>	Sequence 107, APP
2	1949	100.0	391	4	US-09-818-111-102- <del>20</del>	Sequence 102, APP
3	1949	100.0	391	4	US-09-072-596-102- <del>20</del>	Sequence 102, APP
4	1949	100.0	391	4	US-09-072-967-107- <del>20</del>	Sequence 107, APP
5	1949	100.0	391	4	US-09-072-967-107- <del>20</del>	Sequence 107, APP
6	1949	100.0	596	4	US-09-287-849-26- <del>10</del>	Sequence 26, APP
7	1949	100.0	600	4	US-09-287-849-22- <del>10</del>	Sequence 22, APP
8	1944	99.7	729	4	US-09-223-040-2	Sequence 2, APP
9	1944	99.7	729	4	US-09-287-849-2	Sequence 2, APP
10	1652.5	84.8	396	3	US-09-818-112-111	Sequence 111, APP
11	1652.5	84.8	396	4	US-09-818-111-106	Sequence 106, APP
12	1652.5	84.8	396	4	US-09-056-556-111	Sequence 111, APP
13	1652.5	84.8	396	4	US-09-072-596-106	Sequence 106, APP
14	1652.5	84.8	396	4	US-09-072-967-111	Sequence 111, APP
15	1486.5	76.3	359	3	US-09-818-112-109	Sequence 109, APP
16	1486.5	76.3	359	4	US-09-056-556-109	Sequence 109, APP
17	1486.5	76.3	359	4	US-09-056-556-109	Sequence 109, APP
18	1486.5	76.3	359	4	US-09-072-596-104	Sequence 104, APP
19	1486.5	76.3	359	4	US-09-072-667-109	Sequence 109, APP
20	1187.5	60.9	358	4	US-09-287-849-8	Sequence 8, APP
21	1187	60.9	263	3	US-09-818-112-91	Sequence 91, APP
22	1187	60.9	263	4	US-09-818-111-92	Sequence 92, APP
23	1187	60.9	263	4	US-09-056-556-91	Sequence 91, APP
24	1187	60.9	263	4	US-09-072-596-92	Sequence 92, APP
25	1187	60.9	263	4	US-09-072-967-91	Sequence 91, APP
26	766.5	39.3	400	4	US-09-073-019-126	Sequence 126, APP
27	766.5	39.3	400	4	US-09-073-010-126	Sequence 126, APP

## ALIGNMENTS

RESULT 1  
US-09-818-112-107  
; Sequence 107, Application US/08918112  
; Patent No. 6220969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skelley, Yasir A.W.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vecwick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; NUMBER OF SEQUENCES: 153;  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/Ms-DOCS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818-112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Makki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE DOCUMENT NUMBER: 210121:4111C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-6900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-818-112-107

Qy 1 MVDFGALPPEINSARMYAGGSASLVAAGQNWDSVADLSAASAFOSYWGLTVGSWIG 60  
 Db 1 MVDFGALPPEINSARMYAGGSASLVAAGQNWDSVADLSAASAFOSYWGLTVGSWIG 60  
 Qy 61 SSAGIMVAASAPYYAMSVTAGQAELTAACQVRVAAAYETAYGLTVPPIAENRAELMI 120  
 Db 61 SSAGIMVAASAPYYAMSVTAGQAELTAACQVRVAAAYETAYGLTVPPIAENRAELMI 120  
 Qy 121 LIATNLGQNTPAIAVNNEAYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180  
 Db 121 LIATNLGQNTPAIAVNNEAYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180  
 Qy 181 LLEQAAAVERASDAAANQLMNNVPOALQOLOAQPTQGTPSSKLGGILMKTSPHRSPISN 240  
 Db 181 LLEQAAAVERASDAAANQLMNNVPOALQOLOAQPTQGTPSSKLGGILMKTSPHRSPISN 240  
 Qy 241 MVSMANNHNSMTNSCVSMNTLSSMLKGPAAPAAQAVOTIAGNGVRAMSLSLOSSGG 300  
 Db 241 MVSMANNHNSMTNSCVSMNTLSSMLKGPAAPAAQAVOTIAGNGVRAMSLSLOSSGG 300  
 Qy 301 LGGGVAAANLGRAASVGSLSPQAAVAAANGVAPARALPLTSUTSAERGPQMLGGLPV 360  
 Db 301 LGGGVAAANLGRAASVGSLSPQAAVAAANGVAPARALPLTSUTSAERGPQMLGGLPV 360  
 Qy 361 GQMGRAGGGSLGVLRVPPRPMPHSPAAAG 391  
 Db 361 GQMGRAGGGSLGVLRVPPRPMPHSPAAAG 391

## RESULT 2

US-08-818-111-102

; Sequence 102 ; Application US/08818111

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; REED, STEVEN G.

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/818,111

; FILING DATE: 13-MAR-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE DOCKET NUMBER: 210121·417C6

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEX/FAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 102:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 391 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-818-111-102

Query Match 100.0%; Score 1949; DB 4; Length 391;

Best Local Similarity 100.0%; Pred. No. 2.9-154;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPEINSARMYAGGSASLVAAGQNWDSVADLSAASAFOSYWGLTVGSWIG 60

Db 1 MVDFGALPPEINSARMYAGGSASLVAAGQNWDSVADLSAASAFOSYWGLTVGSWIG 60

Qy 61 SSAGIMVAASAPYYAMSVTAGQAELTAACQVRVAAAYETAYGLTVPPIAENRAELMI 120

Db 61 SSAGIMVAASAPYYAMSVTAGQAELTAACQVRVAAAYETAYGLTVPPIAENRAELMI 120

Qy 121 LIATNLGQNTPAIAVNNEAYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180

Db 121 LIATNLGQNTPAIAVNNEAYGENWAQDAAMFGYAAATAATATLIPFEAPEMTSAGG 180

Qy 61 SSAGIMVAASAPYYAMSVTAGQAELTAACQVRVAAAYETAYGLTVPPIAENRAELMI 120

Db 61 SSAGIMVAASAPYYAMSVTAGQAELTAACQVRVAAAYETAYGLTVPPIAENRAELMI 120

RESULT 3

US-08-818-111-107

; Sequence 107 ; Application US/09056556

; Patent No. 6350456

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; REED, STEVEN G.

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Skeiky, Yasir A.W.

; ADDRESS: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; ZIP: 98104-7092

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,556

; FILING DATE: 07-APR-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

US-09-056-556-107

Query Match 100.0% Score 1949; DB 4; Length 391;  
Best Local Similarity 100.0% Pred. No. 2.9e-154;  
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPBEINSARMYAGPGASLVAQQMMDSVASDLSASAFAQSVNGLTVGSWIG 60  
Db 1 MVDFGALPBEINSARMYAGPGASLVAQQMMDSVASDLSASAFAQSVNGLTVGSWIG 60

Qy 61 SSAGLMVAAASPYAMSTTAGOLELTAAQVRAAAAYETAYGLTVPPVILENRALMI 120  
Db 61 SSAGLMVAAASPYAMSTTAGOLELTAAQVRAAAAYETAYGLTVPPVIAENRALMI 120

Qy 121 LIANLLGGNTPAIAVNELEYGENWAQDAAMGYAATAATATLIPPEAPEMTSAG 180  
Db 121 LIANLLGGNTPAIAVNELEYGENWAQDAAMGYAATAATATLIPPEAPEMTSAG 180

Qy 181 LLEQAAAVEBEASDTAAQMLNNPQALQOLAOPTGTTSSKLGGIWKTVSPHRSPISN 240  
Db 181 LLEQAAAVEBEASDTAAQMLNNPQALQOLAOPTGTTSSKLGGIWKTVSPHRSPISN 240

Qy 241 MVSHANNHSMTNNGVSMMTNLSSMLKGAPALAAAQTAONGVRAAMSSGSSG 300  
Db 241 MVSHANNHSMTNNGVSMMTNLSSMLKGAPALAAAQTAONGVRAAMSSGSSG 300

Qy 301 LGGGYAANIGRAASYGSLSYPOQWAIAANGAVTTPAARALPITSLSAERGPQMLGGLPV 360  
Db 301 LGGGYAANIGRAASYGSLSYPOQWAIAANGAVTTPAARALPITSLSAERGPQMLGGLPV 360

Qy 361 QMGARAGGGSLGSLGVLRVPBPYPMPHSPLAG 391  
Db 361 QMGARAGGGSLGSLGVLRVPBPYPMPHSPLAG 391

RESULT 4  
US-09-072-596-102  
Sequence 107, Application US/09072596

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Campos-Neto, Antonio  
 APPLICANT: Houghton, Raymond  
 APPLICANT: Vedula, Thomas S.  
 APPLICANT: Twardzik, Daniel R.  
 APPLICANT: Lodes, Michael J.  
 APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
 NUMBER OF SEQUENCES: 350  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED AND BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.3.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,596  
 FILING DATE: 05-MAY-1998  
 CLASSIFICATION: IBM PC compatible  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.417C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900

RESULT 5  
US-09-072-967-107  
Sequence 107, Application US/09072967

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Campos-Neto, Antonio  
 APPLICANT: Houghton, Raymond  
 APPLICANT: Vedula, Thomas S.  
 APPLICANT: Twardzik, Daniel R.  
 APPLICANT: Lodes, Michael J.  
 APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
 NUMBER OF SEQUENCES: 355  
 CORRESPONDENCE ADDRESS:  
 ZIP: 98104-7032  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.3.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,967  
 FILING DATE: 05-MAY-1998  
 CLASSIFICATION: IBM PC compatible  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.417C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900

FILING DATE: 05-MAY-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.411C9  
 TELEPHONE: (206) 622-4900  
 TELEX/FAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 107:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 391 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: Linear  
 US-09-722-967-107

Query Match 100.0%; Score 1949; DB 4; Length 391;  
 Best Local Similarity 100.0%; Prod. No. 2.9e-154; Mismatches 0; Indels 0; Gaps 0;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPINSARMYAGGSASLVALAQMWDSVASDLSASAFAQSVMWGLTVGSWIG 60  
 Db 1 MVDFGALPPINSARMYAGGSASLVALAQMWDSVASDLSASAFAQSVMWGLTVGSWIG 68

Qy 1 MVDFGALPPINSARMYAGGSASLVALAQMWDSVASDLSASAFAQSVMWGLTVGSWIG 60  
 Db 1 MVDFGALPPINSARMYAGGSASLVALAQMWDSVASDLSASAFAQSVMWGLTVGSWIG 60

Qy 61 SSAGIMVAAASPYYAMSVTAQGELTAQVRYAAAAETAYGLTPPVIAENRAELMI 120  
 Db 61 SSAGIMVAAASPYYAMSVTAQGELTAQVRYAAAAETAYGLTPPVIAENRAELMI 120

Qy 121 LIATNLGQNTPAIVNEAEGYGEWAQDAAMFGYAAATAATATATLIPPEAPEMTSAGG 180  
 Db 121 LIATNLGQNTPAIVNEAEGYGEWAQDAAMFGYAAATAATATLIPPEAPEMTSAGG 180

Qy 181 LLEQAAVEEASDTAAANOLMNNVPAQLOLAQOPTGTTPSSKLGGIWKTSVPHRSPISN 240  
 Db 181 LLEQAAVEEASDTAAANOLMNNVPAQLOLAQOPTGTTPSSKLGGIWKTSVPHRSPISN 240

Qy 241 MVSMANNHSMNTNSGVSMNTLSSMLKGFAAAQAQVOTAQNGVAMSIGSSLOSSG 300  
 Db 241 MVSMANNHSMNTNSGVSMNTLSSMLKGFAAAQAQVOTAQNGVAMSIGSSLOSSG 300

Qy 301 LGGGYANLGRAASTGSLSPQAVAAANAVTAPARALPLTSLSAARGPGLGGIPV 360  
 Db 309 LGGGYANLGRAASVGSLSPQAVAAANQVATPAPARALPLTSLSAARGPGLGGIPV 368

Qy 361 QGMGARGGGLSGVLRVPFRPYMPHSPIAG 391  
 Db 369 QGMGARGGGLSGVLRVPFRPYMPHSPIAG 399

RESULT 6  
 US-09-287-849-26  
 Sequence 26, Application US/09287849  
 Patent No. 6627198

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A. W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Alderson, Mark  
 APPLICANT: Campos-Neto, Antonio  
 APPLICANT: Corixa Corporation  
 TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 FILE REFERENCE: 014058-009020US  
 CURRENT APPLICATION NUMBER: 1999-04-07  
 PRIOR APPLICATION NUMBER: US 08/287,849  
 CURRENT FILING DATE: 1997-03-13  
 PRIOR FILING DATE: 1997-10-01  
 PRIOR APPLICATION NUMBER: US 09/025,197  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 09/056,556  
 PRIOR FILING DATE: 1998-04-07  
 PRIOR APPLICATION NUMBER: US 09/223,040  
 PRIOR FILING DATE: 1998-12-30  
 NUMBER OF SEQ ID NOS: 46

Query Match 100.0%; Score 1949; DB 4; Length 596;  
 Best Local Similarity 100.0%; Prod. No. 5.2e-154; Mismatches 0; Indels 0; Gaps 0;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFGALPPINSARMYAGGSASLVALAQMWDSVASDLSASAFAQSVMWGLTVGSWIG 60  
 Db 9 MVDFGALPPINSARMYAGGSASLVALAQMWDSVASDLSASAFAQSVMWGLTVGSWIG 68

Qy 61 SSAGIMVAAASPYYAMSVTAQGELTAQVRYAAAAETAYGLTPPVIAENRAELMI 120  
 Db 69 SSAGIMVAAASPYYAMSVTAQGELTAQVRYAAAAETAYGLTPPVIAENRAELMI 128

Qy 121 LIATNLGQNTPAIVNEAEGYGEWAQDAAMFGYAAATAATATLIPPEAPEMTSAGG 180  
 Db 129 LIATNLGQNTPAIVNEAEGYGEWAQDAAMFGYAAATAATATLIPPEAPEMTSAGG 188

Qy 181 LLEQAAVEEASDTAAANOLMNNVPAQLOLAQOPTGTTPSSKLGGIWKTSVPHRSPISN 240  
 Db 189 LLEQAAVEEASDTAAANOLMNNVPAQLOLAQOPTGTTPSSKLGGIWKTSVPHRSPISN 248

Qy 241 MVSMANNHSMNTNSGVSMNTLSSMLKGFAAAQAQVOTAQNGVAMSIGSSLOSSG 300  
 Db 249 MVSMANNHSMNTNSGVSMNTLSSMLKGFAAAQAQVOTAQNGVAMSIGSSLOSSG 308

Qy 301 LGGGYANLGRAASTGSLSPQAVAAANAVTAPARALPLTSLSAARGPGLGGIPV 360  
 Db 309 LGGGYANLGRAASVGSLSPQAVAAANQVATPAPARALPLTSLSAARGPGLGGIPV 368

Qy 361 QGMGARGGGLSGVLRVPFRPYMPHSPIAG 391  
 Db 369 QGMGARGGGLSGVLRVPFRPYMPHSPIAG 399

RESULT 7  
 US-09-287-849-22  
 Sequence 22, Application US/09287849  
 Patent No. 6627198

GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Skeiky, Yasir A. W.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Alderson, Mark  
 APPLICANT: Campos-Neto, Antonio  
 APPLICANT: Corixa Corporation  
 TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 FILE REFERENCE: 014058-009020US  
 CURRENT APPLICATION NUMBER: 1999-04-07  
 PRIOR APPLICATION NUMBER: US 08/818,112  
 PRIOR FILING DATE: 1997-03-13  
 PRIOR APPLICATION NUMBER: US 08/942,578  
 PRIOR FILING DATE: 1997-10-01  
 PRIOR APPLICATION NUMBER: US 09/025,197  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 09/056,556  
 PRIOR FILING DATE: 1998-04-07  
 PRIOR APPLICATION NUMBER: US 09/223,040  
 PRIOR FILING DATE: 1998-12-30  
 NUMBER OF SEQ ID NOS: 46

SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 22  
 LENGTH: 600  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence:tri-fusion  
 US-09-287-840-22

Query Match 100.0%; Score 1949; DB 4; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 5 2e-154; Indels 0; Gaps 0;  
 Matches 391; Conservative 0; Mismatches 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSASDLSASAFAQSVMGLTVGSWIG 60  
 Db 9 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSASDLSASAFAQSVMGLTVGSWIG 68

Qy 61 SSAGLMVAAASPYAAMSVTAGQELTAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
 Db 202 SSAGLMVAAASPYAAMSVTAGQELTAQVRVAAAAYETAYGLTPPPVIAENRAELMI 261

Qy 121 LIATNLGGONTPAVNEAEYGENWAQDAAMGYAATAATATLPPFEAPEMTSAGG 180  
 Db 262 LIATNLGGONTPAVNEAEYGENWAQDAAMGYAATAATATLPPFEAPEMTSAGG 321

Qy 181 LLEQAAYEEASDTAAANQMNNTYPOQALQOOLAQPTQTPSSKLGWKTVPSPRSPI SN 240  
 Db 322 LLEQAAYEEASDTAAANQMNNTYPOQALQOOLAQPTQTPSSKLGWKTVPSPRSPI SN 381

Qy 241 MVSMANHMSMTNSGYSMTNLSMILKGFAAAAQAVOTAQNGVRAANSLGSSLGSSG 300  
 Db 382 MVSMANHMSMTNSGYSMTNLSMILKGFAAAAQAVOTAQNGVRAANSLGSSLGSSG 441

Qy 301 LGGGVANLGRAASYGSLSPQAAWAOVTPARALPLTSLTSAERGPQMLGGLPV 360  
 Db 442 LGGGVANLGRAASYGSLSPQAAWAOVTPARALPLTSLTSAERGPQMLGGLPV 501

Qy 361 QMGARGAGGGLSGVLRVPPRPPYMPHSPIAG 391  
 Db 502 QMGARGAGGGLSGVLRVPPRPPYMPHSPIAG 532

RESULT 9  
 US-09-287-849-2  
 ; Sequence 2, Application US/09287849  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; FILE REFERENCE: 014058-00020US  
 ; CURRENT APPLICATION NUMBER: US/09/287-849  
 ; PRIOR APPLICATION NUMBER: US/09/287-849  
 ; PRIOR FILING DATE: 1999-04-07  
 ; PRIOR APPLICATION NUMBER: US/09/287-849  
 ; PRIOR FILING DATE: 1997-03-13  
 ; PRIOR APPLICATION NUMBER: US/09/287-849  
 ; PRIOR FILING DATE: 1997-10-01  
 ; PRIOR APPLICATION NUMBER: US/09/25-197  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US/09/056 556  
 ; PRIOR FILING DATE: 1998-04-07  
 ; PRIOR APPLICATION NUMBER: US/09/223, 040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 729  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion  
 US-09-223-040-2

Query Match 99.7%; Score 1944; DB 4; Length 729;  
 Best Local Similarity 99.7%; Pred. No. 1.8e-153; Indels 0; Gaps 0;  
 Matches 390; Conservative 0; Mismatches 1; InDelS 0; Gaps 0;

Qy 1 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSASDLSASAFAQSVMGLTVGSWIG 60  
 Db 142 MVDFGALPPEINSARMYAGPGSASLVAAQMDVSASDLSASAFAQSVMGLTVGSWIG 201

Qy 61 SSAGLMVAAASPYAAMSVTAGQELTAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
 Db 202 SSAGLMVAAASPYAAMSVTAGQELTAQVRVAAAAYETAYGLTPPPVIAENRAELMI 261

Query Match Score 1652.5; DB 3; Length 396;  
 Best Local Similarity 84.9%; Prd. No. 1.4e-129;  
 Matches 337; Conservative 19; Mismatches 34;  
 Indels 7; Gaps 3;

---

Qy 121 LIATNLGQNTPAIAVNEAYGEMQAADAAMFGYAAATAATATLLPPEAPEMTSAGG 180  
 Db 262 LIATNLGQNTPAIAVNEAYGEMQAADAAMFGYAAATAATATLLPPEAPEMTSAGG 321  
 Qy 181 LLEQAAAEEASDTAAANQLMNMYQALQALQAOQPTGGTTPSSKGGLWKTSPHRSPTSN 240  
 Db 322 LLEQAAAEEASDTAAANQLMNMYQALQALQAOQPTGGTTPSSKGGLWKTSPHRSPTSN 381  
 Qy 241 MVSMANNHMSMTNSGSMTNLSMKGPAAPAAQAVQTAQNGVRAMSSESSLSSGG 300  
 Db 382 MVSMANNHMSMTNSGSMTNLSMKGPAAPAAQAVQTAQNGVRAMSSESSLSSGG 441  
 Qy 301 LGGGYAANLGRAASYGSLSIUPQAAANQVTPAARALPLTSLTSAAERPGGMLGGIPV 360  
 Db 442 LGGGYAANLGRAASYGSLSIUPQAAANQVTPAARALPLTSLTSAAERPGGMLGGIPV 501  
 Qy 361 GQMGRAGGGLSGYLVRPPRPPYMHSPAAAG 391  
 Db 502 GQMGRAGGGLSGYLVRPPRPPYMHSPAAAG 532

RESULT 11  
 US-08-818-112-111  
 / Sequence 111, Application US/08818112  
 / GENERAL INFORMATION:  
 / APPLICANT: Reed, Steven G.  
 / APPLICANT: Skeiky, Yasir A. W.  
 / APPLICANT: Campos-Neto, Antonio  
 / APPLICANT: Dillon, Davin C.  
 / APPLICANT: Houghton, Raymond  
 / APPLICANT: Twardzik, Daniel R.  
 / APPLICANT: Vedwick, Thomas S.  
 / TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
 / NUMBER OF SEQUENCES: 153  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: SEED and BERRY LLP  
 / STREET: 6300 Columbia Center, 701 Fifth Avenue  
 / CITY: Seattle  
 / STATE: Washington  
 / COUNTRY: USA  
 / ZIP: 98104-7092  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: Patent In Release #1.0, Version #1.30  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/818.112  
 / FILING DATE: 13-MAR-1997  
 / CLASSIFICATION: 424  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Maki, David J.  
 / REGISTRATION NUMBER: 31.392  
 / REFERENCE/DOCKET NUMBER: 210121.411C6  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (206) 622-4900  
 / TELEFAX: (206) 682-6031  
 / INFORMATION FOR SEQ ID NO: 111:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 396 amino acids  
 / TYPE: amino acid  
 / STRANDEDNESS:  
 / TOPOLOGY: linear

US-08-818-112-111

Query Match Score 1652.5; DB 3; Length 396;  
 Best Local Similarity 84.9%; Prd. No. 1.4e-129;  
 Matches 337; Conservative 19; Mismatches 34;  
 Indels 7; Gaps 3;

---

Qy 1 VVDFGALPPEINSARMYAGPSASILVAACQMDSVASDLFSAASAFQSVVWGLTTGSWIG 60  
 Db 61 SSAGLMVAAASPVAMMSVTAQAOELTAQVRAAAAYETAYGLTVPPVIAERELMI 120  
 Db 61 SSAGLMVAAASPVAMMSVTAQAOELTAQVRAAAAYETAYGLTVPPVIAERELMI 120  
 Qy 121 LIATNLGQNTPAIAVNEAYGEMQAADAAMFGYAAATAATATLLPPEAPEMTSAGG 180  
 Db 121 LIATNLGQNTPAIAVNEAYGEMQAADAAMFGYAAATAATATLLPPEAPEMTSAGG 180  
 Qy 181 LIEQAAAEEASDTAAANQLMNMYQALQALQAOQPTGGTTPSSKGGLWKTSPHRSPTSN 240  
 Db 181 LIEQAAAEEASDTAAANQLMNMYQALQALQAOQPTGGTTPSSKGGLWKTSPHRSPTSN 240  
 Qy 241 MVSMANNHMSMTNSGSMTNLSMKGPAAPAAQAVQTAQNGVRAMSS---IGSSL 296  
 Db 241 MVSMANNHMSMTNSGSMTNLSMKGPAAPAAQAVQTAQNGVRAMSS---IGSSL 296  
 Qy 241 IIVSMANNHMSMTNSGSMTNLSMKGPAAPAAQAVQTAQNGVRAMSS---IGSSL 296  
 Db 241 IIVSMANNHMSMTNSGSMTNLSMKGPAAPAAQAVQTAQNGVRAMSS---IGSSL 296  
 Qy 297 GSSGLGGVANLGRAASVSLSVSPQAAANQVTPAARALPLTSLTSAAERPGGMLGGIPV 356  
 Db 300 GSSGLGGVANLGRAASVSLSVSPQAAANQVTPAARALPLTSLTSAAERPGGMLGGIPV 359  
 Qy 357 GLPVGOMGARAG--GGLSGYLRVPPRPPYMHSPAAAG 391  
 Db 360 GLPLGOTNSGGFFGGSNALRMPRAYVMRVPRAAG 396

RESULT 11  
 US-08-818-111-106  
 / Sequence 106, Application US/08818111  
 / Patent No. 6336852  
 / GENERAL INFORMATION:  
 / APPLICANT: Reed, Steven G.  
 / APPLICANT: Skeiky, Yasir A. W.  
 / APPLICANT: Dillon, Davin C.  
 / APPLICANT: Campos-Neto, Antonio  
 / APPLICANT: Houghton, Raymond  
 / APPLICANT: Twardzik, Daniel R.  
 / APPLICANT: Vedwick, Thomas S.  
 / TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
 / NUMBER OF SEQUENCES: 148  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: SEED and BERRY LLP  
 / STREET: 6300 Columbia Center, 701 Fifth Avenue  
 / CITY: Seattle  
 / STATE: Washington  
 / COUNTRY: USA  
 / ZIP: 98104-7092  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: Patient In Release #1.0, Version #1.30  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/818.112  
 / FILING DATE: 13-MAR-1997  
 / CLASSIFICATION: 424  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Maki, David J.  
 / REGISTRATION NUMBER: 31.392  
 / REFERENCE/DOCKET NUMBER: 210121.417C6  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (206) 622-4900  
 / TELEFAX: (206) 682-6031  
 / INFORMATION FOR SEQ ID NO: 106:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 396 amino acids  
 / TYPE: amino acid  
 / STRANDEDNESS:  
 / TOPOLOGY: linear

US-08-818-111-106

Query Match Score 1652.5; DB 3; Length 396;  
 Best Local Similarity 84.9%; Prd. No. 1.4e-129;  
 Matches 337; Conservative 19; Mismatches 34;  
 Indels 7; Gaps 3;

---

Qy 1 MVDFGALPPEINSARMYAGPSASILVAACQMDSVASDLFSAASAFQSVVWGLTTGSWIG 60

Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
 Best Local Similarity 84.9%; Pred. No. 1.4e-129;  
 Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;  
 Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

Qy 1 MYDFGALPPPEINSARMYAGPGSASLVAQAAQWDSVASDLFSASAFOSVWGLTVGSWIG 60  
 Db 1 VVDGALPPPEINSARMYAGPGSASLVAQAAQWDSVASDLFSASAFOSVWGLTVGSWIG 60  
 Qy 61 SSAGLMVAAASPYAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
 Db 61 SSAGLMVAAASPYAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
 Db 61 SSAGLMVAAASPYAMSVTAGAELTAAQVRVAAAAYETAYGLTPPPVIAENRAELMI 120  
 Qy 121 LIATNLGQNTPAIAVNEAEYGENWQAQDAAMFGYAATAATATLPEEAPMTSAGG 180  
 Db 121 LIATNLGQNTPAIAVNEAEYGENWQAQDAAMFGYAATAATATLPEEAPMTSAGG 180  
 Qy 181 LLEQAAVEEASDTAAANQLMNVPOALQOLOAOPTGTTSSKLGWKTSPHRSPISN 240  
 Db 181 LLEQAAVEEASDTAAANQLMNVPOALQOLOAOPTKSIMPFLSELWKATSPHLSPLSN 240  
 Qy 241 MVSZANNHMSMTNSGVSMINTLSSMLKGFAAAAQAVTAQONGVRAMSS ---LGSSL 296  
 Db 241 IVSMLNHMSMTNSGVSMINTLSSMLKGFAAAAQAVTAQONGVRAMSS ---LGSSL 296  
 Qy 297 GSSGLGGTAAATLGRAASTGSLSVPOAAVAAANOAVTPAARALPLTSITSAAERGPQOMLG 356  
 Db 300 GSSGLGAGTAANLGRAASTGSLSVPOAAVAAANOAVTPAARALPLTSITSAAQTAQPHMLG 359  
 Qy 357 GLPVGQMGGARAG - GGLSGCVLRVPPRPMVPHSPAAG 391  
 Db 360 GLPLGQLTNSGGGFCGVSNALRMPPRAYMFRVPAAG 396

## RESULT 12

US-09-056-556-111

Sequence 111; Application US/09056556

Patent No. 635056

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A. W.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-1092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: 07-APR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:

LENGTH: 396 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-09-056-556-111

## RESULT 13

US-09-072-596-106

Sequence 106; Application US/09072596

Patent No. 6458365

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A. W.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

TREATMENT

NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/072,596

FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031



COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/818,112  
 FILING DATE: 13-MAR-1997  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.411C6  
 TELECOMMUNICATION INFORMATION:  
 TELEFAX: (206) 682-6900  
 TELEPHONE: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 109:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 359 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-818-112-109

Query Match Similarity 76.3%; Score 1486.5; DB 3; Length 359;  
 Best Local Similarity 84.2%; Pred. No. 7.9e-116; Indels 5; Gaps 2;  
 Matches 303; Conservative 16; Mismatches 36;

Qy	1	MVDFGALPBEINSARMAYAGPSASLVAQQMWDVSASDLSASAFSVWGLTVCSWIG
Db	1	VVDGALPFEINSARMAYAGPSASLVAQKMDVSASDLSASAFSVWGLTVCSWIG
Qy	61	SSAGLMLVAASPYAAMSVTAGOELTTAAQVRVAAAAYETAYGLTVBPPVIAENRAELMI
Db	61	SSAGLMLAAASPYAAMSVTAGQQLTAQVRVAAAAYETAYRLTVBPPVIAENRFLMT
Qy	121	LIAATNLIGGNTPTAIVNEPEYGMWQDAAMMGYAAATAATATATATLIPFEPEPMISAGG
Db	121	LIAATNLIGGNTPTAENQAYSQNWQDQEAMGYAAATAATEALIIPEDPLTNPGG
Qy	181	LLEQAAAVEEASDTAAANOLMNNTVPCALQQAQPTGTTPSSKLGGIWKTVPSPHSPTSN
Db	181	LLEQAVAVVEAIDAAANOLMNNTVPCALQQAQPTGTTPSSKLGGIWKTVPSPHSPTSN
Qy	241	MVSMANNHNSMINTSNSGVSMNTLSSMLKGDPAPAAAQAVQTAAONGVDRAMSS--LGSSL
Db	241	VSSTANNHNSMMGTVGSMMNTLHSMLKGJAP-AAAQAVETAAENGWAMSSLGQCLSSL
Qy	297	GSSGIGGGTGAANLGRAASTGSLSYPOANVAAANOVTPARALPLTSITSAERGPGOMLG
Db	300	GSSGLGAGVAAANLGRAASTGSLSYPOANVAAANOVTPARALPLTSITSAQTAPGHMLG

Search completed: August 25, 2004, 03:14:26

Job time : 36 secs

*This Page Blank (uspto)*